1-1-2014

Predicting Mississippi Curriculum Testing Program, Second Edition performance using the Northwest Evaluation Association Measures of Academic Progress

Mary Coleush

Predicting Mississippi Curriculum Testing Program, Second Edition performance using the Northwest Evaluation Association Measures of Academic Progress

By

Mary Cole-Bush

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Educational Psychology
in the Department of Counseling and Educational Psychology

Mississippi State, Mississippi

August 2014

Predicting Mississippi Curriculum Testing Program, Second Edition performance using the Northwest Evaluation Association Measures of Academic Progress

By

Mary Cole-Bush

Approved:

Carlen Henington
(Major Professor)

Debra L.Prince
(Committee Member)

Carmen D. Reisener
(Committee Member)

Tawny E. McCleon
(Committee Member)

Daniel W. Wong
(Graduate Coordinator)

Richard Blackbourn
Dean
College of Education

Name: Mary Cole-Bush

Date of Degree: August 15, 2014

Institution: Mississippi State University

Major Field: Educational Psychology

Major Professor: Dr. Carlen Henington

Title of Study: Predicting Mississippi Curriculum Testing Program, Second Edition performance using the Northwest Evaluation Association Measures of Academic Progress

Pages in Study: 137

Candidate for Degree of Doctor of Philosophy

The purpose of this study was to determine if the Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) reading and math assessments are a valid predictor of performance on the language arts and mathematics Mississippi Curriculum Test, 2nd Edition (MCT2). Additionally, the researcher sought to determine whether student characteristics of gender, ethnicity, and economically disadvantaged status added statistically to the prediction of MCT2 scores.

The researcher used a correlational research design to answer the research questions that guide this study. Regression analyses were performed using IBM Statistical Package for the Social Sciences (SPSS), version 22. Data were collected from a Southern Mississippi school district. Scores from 676 6th grade students and 659 8th grade students were used in this study.

The results of simple linear regression indicate that NWEA-MAP reading and mathematics assessments are a valid predictor of language arts and mathematics MCT2 scale scores for 6th and 8th grade students. Results of multiple regression indicate that the

linear combination of fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 language arts scale scores for sixth grade students; likewise, the linear combination of fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 language arts scale scores for eighth grade students.

Similarly, multiple regression analyses indicate that the linear combination of fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 mathematics scale scores for sixth grade students; similarly, the linear combination of fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 mathematics scale scores for eighth grade students.

DEDICATION

This work is dedicated to my grandmother, Addie L. Sanders Cole. I am truly blessed to have had you as my grandmother and mother. Thank you for making me believe in myself and thank you for believing in me. It is because of you that I understand the true value of education and hard work.

To my husband, Marvin, thank you for your support and patience. Thank you for taking such good care of our children while I worked on this project nightly and on weekends. Because I knew my children were safe, happy, and very well taken care of I could devote my energy to finishing this project.

To my children, Madison and Marreko, thank you for understanding when mommy had to go to the office to do her "school work." I am proud of both of you and I am very grateful for your patience and understanding.

ACKNOWLEDGEMENTS

I would like to acknowledge my dissertation committee members, Dr. Henington, Dr. Prince, Dr. McCleon, and Dr. Reisner. I am sincerely grateful for your support and guidance. Thank you, Dr. Henington for your various “pep talks” and words of encouragement. Dr. Prince your sincere investment in my completion of this project, words of encouragement and tough love gave me the strength and courage not to abandon this endeavor. I appreciate you all. I could not have finished this project without you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Background

American education has evolved over the years due to significant changes in federal and state education legislation (e.g., No Child Left Behind [NCLB], Common Core Standards, etc.). Today's American educational system is often characterized as competitive (Ravitch, 2010) due to the fact that standardized test scores are seen as the key factors in deciding whether a school provides a quality education for students. Schools are not necessarily being evaluated based upon measures of the quality of teachers and administrators; rather, teacher and administrator quality are being evaluated based upon their students' high stakes testing performance (Ravitch, 2010). The simple fact is, the best schools are defined as having the highest test scores and the worst schools are defined as having the lowest test scores. Thus, school personnel are held professionally accountable for their students' performance rather than on their own performance.

This new way of measuring accountability is fueled by NCLB, also known as Public Law 107-110, presented during President George W. Bush's administration. NCLB is the reauthorization of the Elementary and Secondary Education Act (ESEA), which focused on closing student achievement gaps (i.e, the gap in achievement between non-minority and minority students, male and female students, disadvantaged and

advantaged students) by providing all children with a fair, equal, and significant opportunity to obtain a high quality education (Maleyko & Gawlik, 2011). The provisions of NCLB are intended to close the achievement gap between high and low achieving students, especially the achievement gap between minority and non-minority students (Maleyko & Gawlik, 2011).

Key features of NCLB include the goal that 95 to 100% of students in public schools score proficient or above in reading, math and science by 2014. States must also create their own accountability and standardized assessment systems. Districts and schools are required to meet yearly progress goals, called Adequate Yearly Progress (AYP), or they may face severe penalties (Chapman, 2007).

Provisions of NCLB set challenging accountability guidelines: (a) states are required to detail how they plan to close the achievement gap and ensure all students achieve academic proficiency (b) parents and communities must be informed about state and school progress through annual state and school district report cards; (c) if a school does not make progress, it must offer supplemental services such as free after school tutoring or take corrective action; and (d) after five years, if a school continues to not make AYP, it must make significant changes (U.S. Department of Education, 2002a). Many schools and districts are having difficulties meeting the demands of NCLB. As a result, President Barack Obama's administration proposed more flexible terms, which will provide states more flexibility from provisions of NCLB (U.S. Department of Education, 2011). The ESEA Waiver supports local and state education reform in exchange for stern state agendas to close the achievement gap, promotes rigorous accountability and ensures that all students are on the right track to graduate from high

school being college and career ready (U.S. Department of Education, 2011). Secretary of Education, Arne Duncan, proposed moving away from the labels many schools receive after they do not meet AYP under NCLB. Mr. Duncan proposes that "We should get out of the business of labeling schools as failures and create a new law that is fair and flexible, and focused on the schools and students most at risk" (McNeil, 2011, p. 12).

According to the Mississippi Department of Education (MDE, 2012c), Mississippi submitted the ESEA Flexibility Waiver request to the U.S. Department of Education on February 24, 2012. The request was approved by the U.S. Department of Education on July 19, 2012. As a result of the approved waiver, Mississippi districts and schools are waived from certain NCLB requirements (MDE, 2012c). Specifically, Mississippi schools:

- are not required to meet 100% proficiency by 2014;
- are not labeled in Title I School Improvement for not meeting AYP;
- are not required to provide Supplemental Educational Services (SES) and Public School Choice if in improvement;
- are not limited to spending 21st Century Community Learning Center funds for extended day/year programs (MDE, 2012c).

According to MDE (2012d), Mississippi's School Performance Classification System will not change as a result of the request but the AYP Model will be restructured. Schools are required to:

- meet proficiency expectations called Annual Measurable Objectives (AMOs) at differentiated rates;

- receive federal designations now called “Priority”, “Focus” and “Reward”;
- address all subgroups and implement focused interventions for low performance;
- assess at least 95% of students.

Beginning the 2011-2012 school term, Mississippi schools were required to meet AMO proficiency expectations. Mississippi’s AMO Subgroup results for the 2011-2012 school term are listed below in Table 1.

Table 1

Mississippi’s 2011-2012 Annual Measurable Objective Subgroup Results

Student Groups	Reading/Language Arts	Mathematics
All Students	Not Met	Met
Students with IEPs	Not Met	Not Met
Limited English Proficient	Not Met	Not Met
Economically Disadvantaged	Not Met	Met
Asian	Met	Met
Black	Met	Not Met
Hispanic	Met	Met
Native American	Not Met	Met
White	Not Met	Met

Note. Subgroups with a status of Not Met did not meet proficiency expectations outlined by the 2011-2012 Annual Measurable Objectives.

In an era of high stakes testing and accountability, states are diligently searching for ways to ensure students are learning what they should be. States are employing more consultants, school psychologists, and educational companies to determine whether or not schools are teaching students what they should be learning. Many of these companies and individuals are using computer-adaptive tests, norm-referenced tests, criterion-referenced tests, teacher made tests and other assessments to help identify student strengths and

weaknesses and to help make instructional decisions that will lead to obtaining proficiency or better on state mandated achievement tests (Shapiro, Smith, & Gebhardt, 2012).

Northwest Evaluation Association (NWEA) is a non-profit organization that offers assessment, professional development, and reporting that utilizes data to make decisions about student learning. NWEA has created a computer-adaptive assessment, Measures of Academic Progress (MAP) that is used in many school districts across the United States. Students who take the MAP assessment obtain a scale score called Rasch Unit (RIT; NWEA, 2012d). NWEA publishes scale alignment studies and goal structures to enable educators to use assessment data at the school, district, and state level. NWEA scale alignment studies, also called linking studies, examine the relationship between MAP assessments and state standardized tests used to measure student achievement. The RIT scale score from MAP assessments that corresponds to the various proficiency levels for each subject and for each student grade are identified in linking studies. NWEA goal structures enable educators to connect the content from state standards to each reporting area of a state aligned MAP assessment (NWEA, 2012b).

In the Mississippi Scale Alignment Study, NWEA conducted research to connect the scale of the Mississippi Curriculum Test, Second Edition (MCT2) mathematics and language arts assessments with NWEA-MAP's RIT scale. Specifically, performance-level scores on the RIT scale that would indicate a good chance of success on the MCT2 mathematics and language arts assessments were identified (NWEA, 2011).

In summary, NWEA offers the type of assessments that will allow schools and districts to identify students' academic strengths and weaknesses. Those assessments

have been shown to also provide teachers with specific information pertaining to the standards, benchmarks and objectives the students have learned and are ready to learn next. Because NWEA assessments are aligned to state standards, teachers are able to target specific areas of need to improve standardized achievement test scores. NWEA-MAP is also used to predict students' standardized achievement test performance. Schools use this information to estimate AYP, measure student growth or lack of growth, and to make instructional decisions about teaching and learning.

Statement of the Problem

Public schools across the country are searching for ways to ensure that their students will achieve state-defined levels of proficiency on state mandated assessments. States achieving proficiency will avoid retributions such as loss of funding, state intervention, re-assignment or termination of staff, and even the chartering of public schools (Goertz & Duffy, 2003).

Mississippi students in Grades 3 – 8 take the MCT2 in May of each school term. Although it is favorable for a student to earn an achievement level of proficient or advanced, students are not required to pass the MCT2; however, MCT2 scores are an integral part of a complex formula used to calculate school and district performance in Mississippi (MDE, 2010). Because of the emphasis placed on state mandated standardized achievement assessments, it is important for schools and districts to be able to accurately estimate student standardized test performance.

According to the results published on the Mississippi Assessment and Accountability Reporting System (MAARS) website, a significant number of sixth and eighth grade students do not obtain performance level scores that correspond to proficient

or advanced on the MCT2. Table 2 demonstrates the percentage of sixth and eighth grade students who obtained a performance level of proficient or greater on the language arts and mathematics MCT2 from 2008 through 2012. Specifically, over the past four school terms the highest percentage of sixth grade Mississippi students to achieve a performance level of proficient or greater on the MCT2 language arts assessment was 57% during the 2011-2012 school term. Over the past four school terms the highest percentage of sixth grade Mississippi students to achieve a performance level of proficient or greater on the MCT2 mathematics assessment was 58% during the 2011-2012 school term. Additionally, over the past four school terms the highest percentage of eighth grade Mississippi students to achieve a performance level of proficient or greater on the MCT2 language arts assessment was 55% during the 2011-2012 school term. Over the past four school terms the highest percentage of eighth grade Mississippi students to achieve a performance level of proficient or greater on the MCT2 mathematics assessment was 68% during the 2011-2012 school term.

Table 2

Percentage of Mississippi Sixth and Eighth Grade Students Scoring Proficient or Advanced on MCT2 Language Arts and Mathematics Assessments

		% Proficient or Advanced			
Grade	Assessment	2011-2012	2010-2011	2009-2010	2008-2009
6	MCT2 Language Arts	57%	54%	53%	51%
6	MCT2 Mathematics	58%	55%	56%	53%
8	MCT2 Language Arts	55%	51%	51%	48%
8	MCT2 Mathematics	68%	66%	60%	54%

Although MCT2 results provide a wealth of diagnostic assessment information to Mississippi teachers, the assessment is given only at the end of the year; therefore, teachers cannot use the information to guide instruction for their students in a timely manner. Providing diagnostic assessment information to the teacher or student before students are required to take and pass a high-stakes assessment affords students the opportunity to receive specific information about their own skills and teachers the opportunity to modify or tailor instruction based on detailed diagnostic assessment results (Sloane & Kelly, 2003). In addition, if Mississippi teachers are able to identify students' strengths and weaknesses at the start of the school term, teachers and schools are able to acquire and provide the necessary resources needed to remediate and strengthen deficits so that a student will have a better opportunity with high stakes assessments.

Because of accountability statutes and the emphasis placed on students' state mandated standardized achievement test scores, it is important that teachers, schools, and districts understand their students' strengths and weaknesses as early as possible. MCT2 results are used in determining Achievement, Growth, and AYP. It is also an integral part of determining whether Mississippi schools adhere to NCLB legislation. Thus, finding an assessment that provides instructional guidance before students take the MCT2 is crucial to obtaining diagnostic information that may assist in improving student outcomes.

Purpose of the Study

Education in the United States has been constantly evolving for more than three decades. Key politicians and stakeholders have invested time and energy into developing rigorous educational practices that will help the United States provide premier educational experiences to its students. Academic achievement standards, accountability

systems, and assessment systems are a few components that play a significant role in our educational evolution. A key factor in determining educational success or failure is student outcomes as measured by state mandated standardized achievement assessment results. Students' standardized assessment results are used to assess student learning, thus school and district efficacy (Goertz & Duffy, 2003).

NCLB imposes stern accountability rules that require states to create their own assessment and accountability systems. Schools and districts are required to meet the demands of their state's assessment and accountability model. Mississippi's accountability model integrates students' standardized achievement scores into a complex formula that assigns accountability ratings at the school and district level (MDE, 2010). Schools and districts that do not achieve favorable ratings could face severe penalties, while schools that do achieve favorable ratings may receive awards and accolades (Flynn, 2008).

It is important for schools and districts to be able to assess what students know and what students are ready to learn next. Utilizing a diagnostic and prescriptive assessment can help schools and districts assess teacher effectiveness. The diagnostic and prescriptive features of such an assessment can be useful at the student, class, or school level by providing educators with specific information about student learning (Neil, 2006).

For the purposes of this study, the researcher examined NWEA-MAP and MCT2 results of students in Grades 6 and 8. Sixth grade students who have attended a Mississippi public school since their third grade year will have taken an assessment within the Mississippi Statewide Accountability System for three consecutive years,

while many eighth grade students who have attended a Mississippi school since grade 3 will have taken an assessment within the Mississippi Statewide Accountability System up to five times. Eighth grade students who pass to the ninth grade and attend a Mississippi high school will be expected to take and pass the assessments that encompass the Mississippi Subject Area Testing Program, Second Edition (SATP2); those assessments are Biology I, Algebra I, English II (writing and multiple choice questions) and United States History. Although the SATP2 will not be a focus in this study, it is important for students who attend schools that participate in the Mississippi Statewide Accountability System to be prepared to take the SATP2 assessments because high school students are expected to take and pass all assessments within SATP2 before graduating from high school. It is also important for eighth grade students to be prepared to enter high school and prepared to encounter the next level of standardized assessments. Eighth grade students' MCT2 results can be seen as an indicator of whether or not those students are adequately prepared to enter high school and to begin taking assessments within Mississippi's SATP2. Moreover, sixth grade students' MCT2 results can be seen as an indicator of whether or not those students are adequately prepared to master seventh grade content. It is important for both sixth and eighth grade students to meet proficiency expectations because Mississippi schools are required to meet AMO goals.

The purpose of this study is to determine whether the reading and mathematics NWEA-MAP assessments, as well as, characteristics of gender, ethnicity, and economically disadvantaged status predict sixth and eighth grade students' performance on the language arts and mathematics MCT2. Predicting MCT2 performance will assist teachers, schools, and districts with making data-driven decisions about improving

instruction and providing students with prescriptive information relating to increasing their academic achievement.

Research Questions

1. How accurately do fall NWEA-MAP RIT reading scores predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade?
2. How accurately do fall NWEA-MAP RIT mathematics scores predict mathematics achievement scores as measured by the MCT2 mathematics scale scores for students in sixth and eighth grade?
3. How accurately do spring NWEA-MAP RIT reading scores predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade?
4. How accurately do spring NWEA-MAP RIT mathematics scores predict mathematics achievement scores as measured by the MCT2 mathematics scale scores for students in sixth and eighth grade?
5. How accurately do the fall and spring NWEA-MAP RIT reading scores, gender, ethnicity and economically disadvantaged status predict student performance on the MCT2 language arts assessment for students in sixth and eighth grade?
6. How accurately do the fall and spring NWEA-MAP RIT mathematics scores, gender, ethnicity and economically disadvantaged status predict student performance on the MCT2 mathematics assessment for students in sixth and eighth grade?

Significance of the Study and Summary

The implications of high-stakes testing and accountably necessitate states, districts, and schools to find ways to prepare students for successful achievement on mandated state standardized assessments. Therefore, there is a need for research surrounding the use of methods and assessments proven to provide accurate information about student growth and achievement so that schools may use this information to remediate students who are at-risk of failing state mandated standardized achievement assessments (Lee, 2007).

Implications of high-stakes testing also necessitate school psychologists to be prepared to identify and assess barriers to academic achievement. According to the National Association of School Psychologists (NASP), a part of a school psychologist's duties includes working with students, parents, teachers and administrators to increase academic achievement. Specifically, a major part of school psychologists' responsibilities includes conducting assessments to identify a student's academic deficits and making instructional recommendations to teachers, parents, and administrators to improve student outcomes (NASP, n.d.). If NWEA-MAP assessments provide diagnostic academic information, school psychologists may be able to use student results to help identify and address academic concerns.

A review of relevant literature about academic standards, NCLB, high-stakes testing, computer-adaptive assessments, NWEA-MAP, Mississippi Statewide Accountability System and MCT2 will be presented in Chapter Two. The presentation and progression of the materials will allow the reader to understand how the subjects are interconnected and have led to the current era of educational practices. In Chapter Three,

research methodologies will be discussed. Chapter Four will contain the summary and results, and Chapter Five will contain the findings and conclusions.

CHAPTER II

LITERATURE REVIEW

Introduction

The purpose of the current study is to determine whether sixth and eighth grade students' language arts and mathematics MCT2 achievement scores can be predicted by their NWEA-MAP reading and mathematics RIT scores. Mississippi students in public schools are required to take state standardized achievement assessments. The creation of educational standards is the precursor to high-stakes testing. The current chapter provides the reader with an overview of important literature related to the history of the standards movement and high-stakes testing. An overview and discussion of NCLB, computer-adaptive assessments, NWEA-MAP, and MCT2 are also offered.

Definition and Purpose of Academic Standards

According to the U.S. Department of Education (2002b), academic standards, sometimes called education standards, should detail what students are expected to know and be able to do, contain coherent and rigorous content and encourage the teaching of advanced skills. A review of the history of American education reveals that many political leaders believed there were various obstacles that limited students from leading the nation educationally (National Commission on Excellence in Education, 1983). Those limitations were also visible in the workforce; thus, affecting the technological

advancement of the United States and its ability to compete with other countries (National Commission on Excellence in Education, 1983). The creation of academic standards began the movement to improve the lives of Americans economically and educationally. The purpose of academic standards is to provide a foundation upon which states can build educational systems that provide students with a rigorous, challenging education that prepares them for our technologically advancing society.

History of Standards Movement and High-Stakes Assessment

Concerns about failing educational systems can be traced back more than three decades. During President Ronald Reagan's administration, it was observed that American schools were not producing young adults who were ready for a competitive workforce. Moreover, American students were not faring well academically when compared to students in other countries. As a result, in 1981, then Secretary of Education T. H. Bell, created the National Commission on Excellence in Education. He charged this commission with identifying the problems with American education and presenting solutions (National Commission on Excellence in Education, 1983).

The National Commission on Excellence in Education consisted of 18 members chosen by the Secretary of Education. The Commission was to review various aspects of American education and report back to the Secretary of Education within 18 months. The report submitted by the Commission was titled *A Nation at Risk: The Imperative for Educational Reform*. In this report, the members summarized educational statistics, various observations, and highlighted several reasons why they felt the nation was at risk. Among those reasons was the fact that during the 1970s, American students did not score amongst the top percentile on 19 academic tests, in fact, they were last on eight

occasions; during that time, more than 23 million American adults were considered functionally illiterate based on tests of reading, writing, and comprehension. Seventeen percent of all 17-year olds in the United States were considered functionally illiterate. Scholastic Aptitude Tests (SAT) results revealed that there was a steady decline in the scores of American students from 1963 to 1980. Furthermore, business and military leaders reported having to spend a considerable amount of money on training programs and educating individuals in basic reading, writing, spelling, and computation (National Commission on Excellence in Education, 1983).

Our nation was faced with the aforementioned challenges and many more during an era of technological advancements, including transformations in health care, medical science, energy production, food processing, and construction. Government officials sought to find solutions so that the United States could compete with increasing technological advances. The primary focus of the Commission was to identify weak areas in American education and offer solutions to strengthen educational structures. According to government officials, this task was integral in order to compete with other nations (National Commission on Excellence in Education, 1983).

The publication of *A Nation at Risk: The Imperative for Educational Reform* was an integral part of the standards movement. Soon after the report, a series of efforts to reform education emerged. In September 1989, President George H. W. Bush assembled the first National Education Summit. The purpose of the Summit was to collaborate about strategies that would help to strengthen the American education system in order to ensure the nation's workforce would be adequately prepared with the knowledge and skills needed to compete in an increasingly comprehensive economy. Like the National

Commission on Excellence in Education, the National Education Summit recognized the need for rigorous education standards. The National Education Summit emphasized the development of standards for student performance and adopted a set of National Education Goals. The goals, targeted for the year 2000, laid the foundation for education improvement at all stages of an individual's life (National Education Goals Panel, 1999).

Shortly after the 1989 National Education Summit, President George H. W. Bush began his tenure as the 41st United States President. In his 1990 State of the Union Address he discussed education reform.

> By the year 2000, every child must start school ready to learn. The United States must increase the high school graduation rate to no less than 90 percent. And we are going to make sure our schools' diplomas mean something. In critical subjects -- at the 4th, 8th, and 12th grades -- we must assess our students' performance. By the year 2000, American students must be first in the world in math and science achievement. Every American adult must be a skilled, literate worker and citizen. Every school must offer the kind of disciplined environment that makes it possible for our kids to learn. And every school in America must be drug-free. Ambitious aims? Of course. Easy to do? Far from it. But the future's at stake. The Nation will not accept anything less than excellence in education. These investments will keep America competitive. And I know this about the American people: We welcome competition. We'll match our ingenuity, our energy, our experience and technology, our spirit and enterprise against anyone. But let the competition be free, but let it also be fair. America is ready (Bush, 2009, pp. 6-7).

President George H. W. Bush's speech verbalized his belief that meeting education goals will keep America competitive with other countries during a time of technological and global advancements. In February of 1990, President George H. W. Bush announced the National Education Goals (National Education Goals Panel, 2002).

In 1994, President William Clinton signed the Goals 2000: Educate America Act. This act provided financial resources to states and communities that developed and submitted school improvement plans designed to ensure all students reach their fullest potential. The Goals 2000: Educate America Act is known as a standards-based education reform because of its focus on student outcomes as the determinant for meeting many of the goals set forth in the act. In the Goals 2000: Reforming Education to Improve Student Achievement report, Goals 2000 is attributed to helping 36 states establish content standards in core academic areas and helping 17 states and Puerto Rico establish performance standards; the remaining states were developing standards. Moreover, states were developing assessments that were aligned with their standards and were expected to have them completed by 2001; in addition to aligned assessments, states were also developing accountability measures and enhancing teacher education programs and professional development efforts to support the standards and accountability movement (U.S. Department of Education, 1998).

In summary, the history of American education reveals many concerns about our failing educational systems. There were several attempts to address the issue. The National Commission on Excellence was one of the first attempts to diagnose our failing educational systems. Their findings led to the creation of educational standards. Since

that time, our educational standards have been refined, creating a more standardized way to measure academic achievement.

Common Core State Standards

In the years to come, the efforts to continue the standards movement and improve education in the United States have continued. Common Core State Standards are the most recent set of education standards introduced in the United States (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Common Core State Standards for English Language Arts and Mathematics were developed (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). To date, 45 states, the District of Columbia, four territories, and the Department of Defense Education Activity have adopted Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

Mississippi adopted Common Core State Standards in 2010 (MDE, 2013a). Although Mississippi has adopted Common Core State Standards, the MCT2 is not aligned with Common Core State Standards; instead, items on the language arts and mathematics MCT2 are aligned with the 2006 Mississippi Language Arts Framework – Revised and the 2007 Mississippi Mathematics Framework – Revised. The researcher will not present detailed information about Common Core State Standards; rather, detailed information about the 2006 Mississippi Language Arts Framework – Revised and the 2007 Mississippi Mathematics Framework – Revised will be presented in the literature review.

No Child Left Behind

In an era of accountability, the standards movement has continued with policy makers searching for ways to ensure all students receive a high quality education. NCLB continued the movement to make American education better. NCLB, also known as Public Law 107-110, was signed into law by President George W. Bush on January 8, 2002. NCLB is the reauthorization of ESEA. NCLB includes Title I, a program that provides funding for local education agencies and schools that educate large numbers of economically disadvantaged students to ensure that those students receive the support necessary to meet state standards (U.S. Department of Education, 2002a). NCLB expanded the federal government's role in education. It was created to ensure states, districts, and schools were held accountable for student achievement, especially disadvantaged students.

Title I of NCLB is intended to ensure that all students have an equal opportunity to obtain a high-quality education as evidenced by obtaining standardized test scores at a level of proficient (U.S. Department of Education, 2002a). According to the mandate, states can achieve this purpose by ensuring several efforts are in effect. Title I proposes that states make sure schools are using high-quality assessments and accountability systems, ensure teachers receive high-quality preparation and training, and ensure schools use curriculum and instructional materials that are aligned with rigorous state academic standards. NCLB also requires states to address the needs of low achieving students in high poverty schools, limited English proficient, migrant, special education, Native American, neglected or delinquent, and young children who demonstrate reading deficiencies (Buchsbaum, 2013). States must close the achievement gap between high

and low performing students, specifically the gap between minority and nonminority students and advantaged and disadvantaged students (Maleyko & Gawlik, 2011). Achievement gap can be defined as the difference in achievement between Caucasian and minority students (Carpenter & Ramirez, 2012). Research has shown that several factors may contribute to the achievement gap of different groups of students. Socioeconomic status has been cited as having influence on student achievement (Tajalli & Opheim, 2004). Gibb, Fergusson, and Horwood (2008) assert that gender also influences student achievement, with females performing better than males.

Title I proposes that states, local education agencies, and schools are responsible for improving the academic achievement of all students. Resources should be allotted appropriately to ensure that areas with the greatest need receive adequate support. NCLB provisions ascertain that using state assessment systems, schools should improve and strengthen accountability, teaching and learning. States will provide schools with more autonomy in exchange for more responsibility for student performance. States will provide students with accelerated programs and programs that increase quality instruction; provide students with access to scientifically based instructional strategies and rigorous academic content. Title I proposes that states improve the quality of education by providing staff with professional development, coordinating services with other agencies to provide programming to students and their families, and enabling parents to participate in educating their children (U.S. Department of Education, 2002a). NCLB also introduced Reading First, a federal program designed to encourage using scientifically based reading research to determine the best methods and assessments to use in early reading instruction (U.S. Department of Education, 2002a).

NCLB has set rigorous guidelines for teachers. According to Smith and Gorard (2007), teachers of every core academic subject must be highly qualified. In order for a teacher to obtain the highly qualified status, he or she must earn full certification or pass the state's teacher certification test for the subject they wish to teach; this includes alternative certification programs. In addition to the aforementioned provisions, NCLB requires states to publish annual report cards that outline student achievement at the state, district, and school level (U.S. Department of Education, 2002a).

The Effects of NCLB

According to Harriman (2005), one of the major purposes of NCLB is to close the achievement gap between advantaged and disadvantaged students; disadvantaged students include students with disabilities, minorities, English language learners, and students in small rural communities. The provisions set forth in NCLB to ensure that the achievement gap is closed, elicits mixed actions and reactions from states, districts, schools, educators, parents, and students. Many reactions have been negative. Some feel the mandates set forth by NCLB promote negativity and anxiety for students and educators (Harriman, 2005). NCLB provisions mandate significant changes in many aspects of education, including school choice, teacher credentials and the way we measure student learning (Harriman, 2005).

NCLB has induced an array of perceptions from students and educators. Students, who are made aware of the stipulations set forth by NCLB, acknowledge an increased stress level brought on by increased emphasis on testing (Harriman, 2005). Although students recognized the advantage of school choice, there was a level of anxiety attached to moving to a new school if their school is failing, "I think students will try harder so

they won't have to trade schools" (Harriman, 2005, p. 66). When asked about AYP rankings, the students recognized that the process may be unfair to schools because of the different individual needs of students, "… kids don't learn at the same rate so they shouldn't hold the school accountable" (Harriman, 2005, p. 66). Although educators acknowledge that NCLB's strict accountability guidelines causes teachers to consider how, what, and why they are teaching more thoroughly, the negative effects cannot go unmentioned. According to Harriman (2005), educators expressed concern about the bureaucratic nature of NCLB, as well as the paperwork requirements.

In an effort to meet the demands of NCLB, school administrators have restructured the traditional school system. Instructional programs, recess, music, art, gifted programming, programming for exceptional children, as well as other elective courses have been eliminated or altered in order to increase programming that is expected to improve standardized test scores (Henley, McBride, Milligan, & Nichols, 2007). According to Pederson (2007), state assessment directors acknowledged that resources and time were reduced for non-assessed subjects, as well as, an increase of curriculum integration of non-assessed subjects with the assessed subjects; there was also an increase in the alignment of curriculum and assessment with state standards for non-assessed subjects. NCLB has also been found to have an indirect effect on Arts Education. Teacher interviews indicate changes in the curriculum and decreases in arts learning opportunities (Spohn, 2008). These decisions, like many other were administrative decisions that were intended to improve test scores and meet the demands of NCLB.

High-Stakes Testing

The publication of *A Nation at Risk* began the movement of improving the educational outcomes of American students by creating rigorous academic standards and accountability measures (National Commission on Excellence in Education, 1983). The evolution of high-stakes assessments aligned with rigorous academic standards are integral parts of state accountability systems (Shields, 2008). NCLB requires states to have stricter accountability rules for local education agencies. This means students are tested more, schools have more rigorous improvement guidelines, and greater sanctions are imposed for schools and districts that do not meet their accountability goals. Although testing has been a part of educating students for quite some time, NCLB has placed a significant focus on testing, making it an integral determinant of school success. According to Goertz and Duffy (2003), testing has a significantly new meaning for students, teachers and schools as a result of the focus on academic standards and accountability.

Fraenkel, Wallen, and Hyun (2012) define high-stakes testing as using tests as the primary, or only, basis for decisions having major consequences. Of note, those consequences may vary in degree (positive or negative) or severity for different stakeholders. For students, consequences could include grade retention, grade promotion or even denial of a high school diploma. For teachers or administrators, consequences could include a change in job placement, demotion, termination or promotion and even merit pay. For schools and districts, negative consequences could include public condemnation, sanctions, or closure; positive consequences could include public praise, awards or funding. According to Plake (2002) high-stakes testing significantly effects

decisions that have important and long-term effects on students. High-stakes tests are used to identify students who need remediation, make retention decisions, and dictate high school graduation status.

Sloane and Kelly (2003) also note other possible positive and negative effects of high-stakes assessments for students. Some positive effects of high-stakes testing for students are: (a) students are provided with clearer information about their own knowledge and skills, (b) students are motivated to work harder in school, (c) high-stakes assessments send clearer signals to students about what to study, and (d) high-stakes assessments help students associate and align personal effort with rewards. Potential negative effects of high-stakes testing for students include: (a) students may become frustrated or discouraged from trying to do their best on the assessment, (b) students may become more competitive, and (c) students may begin to devalue grades and school assessments (Sloane & Kelly, 2003). Moreover, some students may experience a lack of confidence or low morale when it comes to high-stakes testing (Shields, 2008).

Although high-stakes testing was designed to create positive outcomes for student achievement, in some instances the negative consequences overshadow the positive benefits (Faulkner & Cook, 2006). High-stakes testing has potential benefits in that it provides a way for the community to see academic progress or lack of progress, provides a focus for the curriculum, enables schools to set performance goals, and schools and districts receive federal funding based on test scores (Faulkner & Cook, 2006). Some of the negative consequences of high-stakes testing include teacher, administrator, and student cheating, exclusion of low performing students from testing, misrepresentation of student dropouts, teaching to the test, narrowing the curriculum, conflicting

accountability ratings, questions about the meaning of proficiency, declining teacher morale, and score reporting errors (Faulkner & Cook, 2006). Also, there is no direct relationship to high-stakes testing and improving student academic performance (Faulkner & Cook, 2006). According to Assaf (2008), other negative effects of high-stakes testing include students experiencing anxiety, fear, aggression, low motivation and low self-esteem, teachers deciding to leave low performing schools, teachers deciding to change the grade levels they teach or teachers deciding to leave teaching all together. Because of the pressure and consequences associated with high-stakes testing, teachers are presenting information to students in the most efficient ways, rarely making higher level connections (Vogler & Virtue, 2007). Although there are many unintended negative consequences to high-stakes testing, without high-stakes tests, many low-performing students and schools would go unnoticed and not get the additional resources needed to improve student and school outcomes.

Sloane and Kelley (2003) assert that there have been different responses to the use of high-stakes testing in education. The notion that a state or federal government can combine a student's high-stakes assessment score with other students high-stakes assessment scores and declare a school as failing provides minimal support to a student or school without remedial resources (Sloane & Kelly, 2003).

> A single test score cannot possibly measure a school's effectiveness, yet schools' funding and reputations rest on test results – perhaps because scores are easy for voters to understand. In reality, voters alike should be looking at multiple measures of a school's effectiveness (Edwards & Pula, 2011, p. 13).

Because high-stakes assessments are given at the end of the school year, these assessments rarely provide diagnostic information for the student or teacher and the information gained is not received in a timely manner to be useful to remediate the student for the current school year (Sloane & Kelly, 2003). According to the American Educational Research Association (2000), no test is valid for all purposes; therefore, when using a high-stakes assessment to improve student and school outcomes, it is very important to make sure the test is sound, scored properly, and used appropriately. The American Psychological Association, American Educational Research Association, and National Council on Measurement in Education created the Standards for Educational and Psychological Testing, where they present principles that are intended to encourage fairness in testing and avoid inadvertent consequences (American Educational Research Association, 2000).

In the plight of increased accountability, more emphasis has been placed on state mandated standardized achievement assessments, often referred to as high-stakes assessments. According to Smyth (2008), the American education system has become heavily reliant on test scores. States are holding school districts accountable for a not obtaining proficiency on state mandated achievement assessments. As a result, more teachers are faced with the decision of providing a traditional, exploratory learning environment for students or "teaching to the test through drill and kill" (Smyth, 2008, p. 134).

High-stakes assessments are assessments of learning rather than assessments for learning (Sloane & Kelly, 2003). According to Neill (2006), there is a need for high-quality assessments that can contribute to improving student achievement by providing

specific information for school improvement. Neill also notes that there is one major limit of testing; state standards are typically "too long and detailed to ever be completely taught to students" (Neill, 2006, p. 9). Neill also explains that assessments that allow teachers to identify educational strengths and weaknesses are useful at all levels, individual, classroom, school, and district. Having this information will allow educators to make informed decisions regarding teaching and learning.

High-stakes testing is a requirement of NCLB. Students' high-stake test scores are used to make decisions about the curriculum and instruction, as well as, teacher, student, school and district achievement. Although the movement to place more emphasis on student test scores was intended to produce positive outcomes for students, there have been many notable negative effects (Sloane & Kelly, 2003).

National Assessment of Educational Progress

In an era of high-stakes testing, it is integral that our educational system has a subjective process for measuring student achievement nationally. Because each state is responsible for creating its own accountability and assessment system, using state-mandated assessment results may not be the most reliable source. The National Assessment of Educational Progress (NAEP) is a nationally representative assessment that may be used for this purpose.

According to the National Center for Education Statistics (NCES), NAEP is an ongoing assessment that is comprised of the largest representative sample of student who are assessed in the areas of mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history; starting in 2014, students will be tested in Technology and Engineering Literacy (NCES, 2012). The NAEP assessment is

administered across the nation using the same test administration procedures, with the test remaining the same year after year, with the exception of well documented changes; as a result, NAEP assessment results provide national comparisons of student academic progress over time (NCES, 2012). According to the NCES (2012), NAEP does not provide individual student or school scores; on the contrary, NAEP results are based on the results of a representative sample of students in grades 4, 8, and 12 for the main assessments or students ages 9, 13, or 17 for long-term trend assessments.

MDE reports NAEP assessment results using the MAARS database. Tables 3 and 4 represent Mississippi's fourth and eighth grade students' 2011 NAEP Language Arts and Mathematics results in comparison to the national representative sample. Based on the information provided, the percent of fourth and eighth grade students who scored proficient or above on the NAEP Language Arts assessment in Mississippi did not vary from the national sample. However, the percent of Mississippi students who scored basic or above on the NAEP Language Arts assessment was lower than the national sample. Additionally, fourth and eighth grade NAEP Language Arts Mississippi mean scale scores were lower than the United States mean scale scores in fourth and eighth grade (MDE, 2013b).

Additionally, the percent of fourth and eighth grade students in Mississippi who scored proficient or above on the NAEP Mathematics assessment was lower than the national sample. Moreover, the percent of Mississippi students who scored basic or above on the NAEP Mathematics assessment was lower than the national sample. Additionally, Mississippi fourth and eighth grade students' NAEP Mathematics mean scale scores were lower than the United States mean scale scores in fourth and eighth grade (MDE, 2013b).

Table 3

NAEP 2011 Language Arts Results

	Mean Scale Score		Percentage At or Above Basic		Percentage At or Above Proficient	
Grade	MS	U.S.	MS	U.S.	MS	U.S.
4	209	220	55	66	22	32
8	254	264	65	75	21	32

Note. The information from this table was taken from the Mississippi Report Card.

Table 4

NAEP 2011 Mathematics Results

	Mean Scale Score		Percentage At or Above Basic		Percentage At or Above Proficient	
Grade	MS	U.S.	MS	U.S.	MS	U.S.
4	230	240	72	81	25	39
8	269	283	58	73	19	34

Note. The information from this table was taken from the Mississippi Report Card.

In summary, NAEP assessments provide states with a standardized way to measure the achievement of its students. NAEP assessments also provide states, districts, and schools with reliable assessment data they may use to determine how well their students are performing compared to other students across the nation. NAEP assessments do not provide individual or school scores; therefore, schools must find alternate ways to obtain more specific assessment information about its students.

Computer-Adaptive Tests

The evolution of education, particularly the shift to more rigorous standards, high-stakes testing, and accountability has led to educators using computer-adaptive tests. Computer-adaptive testing emerged in K-12 education over two decades ago (Kingsbury

& Hauser, 2004). According to NWEA (2012e), computer-adaptive tests are administered on the computer and the level of difficulty is adjusted based on a student's performance; if a student answers a question correctly, the difficulty of the questions increases and if a student answers a question incorrectly the questions become easier. This method allows a student to be tested at his or her individual instructional level. Because students are tested at their instructional level, computer-adaptive assessments provide diagnostic information specific to instruction (Shapiro & Gebhardt, 2012). Computer-adaptive testing has been referenced as an assessment for learning rather than an assessment of learning; on the contrary, high-stakes testing is generally referenced as an assessment of learning (Sloane & Kelly, 2003).

The foundation of computer-adaptive assessments ascertains that "there is a progression of skills underlying the academic domain being assessed" (Shapiro & Gebhardt, 2012, p. 296). The item-response theory is used when creating computer-adaptive assessments for "adapting the item difficulty level to the person's knowledge level" (Wauters, Desmet, & Van den Noortgate, 2010, p. 550).

According to Delong (2007), computer-adaptive assessments may be beneficial in that they enable educators to test students on local or state standards in an efficient and practical manner. Computer-adaptive assessments are being used in K-12 education for screening and diagnostic assessment (Shapiro & Gebhardt, 2012). A few computer-adaptive assessments that are being used in K-12 education are STAR Reading, STAR Reading Spanish, STAR Math, STAR Early Literacy, MAP Reading, MAP Mathematics, MAP Science, MAP Language and MAP for Primary Grades (reading and math).

STAR Reading, STAR Math, STAR Early Literacy, and STAR Reading Spanish are computer-adaptive assessments created by Renaissance Learning (Renaissance Learning, 2011). STAR Reading, also available in Spanish, is a screening and progress monitoring assessment that provides a measure of general reading achievement and comprehension; it is designed for students in Grades 1 through 12 (Renaissance Learning, 2011). STAR Math is a diagnostic, screening and progress monitoring assessment that provides a measure of general math achievement; it is designed for students in Grades 1 through 12 (Renaissance Learning, 2011). STAR Early Literacy is a computer-adaptive screening, progress monitoring, and diagnostic assessment that assesses 41 skills in seven early literacy domains; it is designed for students in Grades K-3 (Renaissance Learning, 2011).

According to Renaissance Learning (2011), STAR Reading and Math are reliable, valid, and provide nationally norm-referenced reading and math scores. Reliability for STAR Reading and Math were assessed using a variety of methods (e.g., split-half, generic, and test-retest). Additionally, validity studies were conducted in schools across 48 states and the District of Columbia linking STAR Reading and Math to state standardized assessments (Renaissance Learning, 2011). Table 5 demonstrates a summary of STAR Reading validity studies. Table 6 demonstrates a summary of STAR Math validity studies.

Table 5

Summary of STAR Reading Validity Studies

	Predictive			Concurrent and Other External Validity		
Grade	Studies	Students	Average Correlation	Studies	Students	Average Correlation
1	6	74,877	.68	15	1,135	.77
2	10	184,434	.78	32	4,142	.72
3	30	200,929	.80	44	4,051	.75
4	25	185,528	.82	41	5,409	.75
5	29	126,029	.82	40	3,588	.75
6	23	82,189	.82	37	2,728	.71
7	23	64,978	.81	33	3,294	.70
8	25	34,764	.81	29	2,148	.72
9	8	9,567	.83	15	949	.72
10	9	7,021	.85	11	566	.61
11	6	6,653	.86	6	324	.70
12	2	3,107	.86	4	165	.74

Table 6

Summary of STAR Math Validity Studies

	Predictive			Concurrent and Other External Validity		
Grade	Studies	Students	Average Correlation	Studies	Students	Average Correlation
1	6	11,880	.55	6	179	.58
2	10	33,076	.63	17	987	.61
3	30	52,604	.66	49	6,400	.61
4	23	55,285	.69	49	5,823	.59
5	29	39,869	.70	58	6,873	.64
6	13	27,663	.73	37	4,202	.66
7	15	18,919	.75	29	3,361	.64
8	11	12,780	.76	29	3,713	.65
9	6	2,545	.78	13	665	.57
10	6	2,236	.79	10	334	.60
11	6	1,921	.80	10	495	.68
12	2	885	.77	9	233	.68

MAP Reading, MAP Mathematics, MAP Science, MAP Language and MAP for Primary Grades (reading and math) are other computer-adaptive assessments that are often used in K-12 education. Because MAP assessments are a focus of this study, an extensive overview of MAP assessments is provided in the next section.

Northwest Evaluation Association and Measures of Academic Progress

NWEA was founded in 1974. The primary goal of NWEA is to improve educational outcomes for students. To this end, NWEA created one of the first computer-adaptive tests (NWEA, 2012a). NWEA computer-adaptive assessments are student centered and provide schools with accurate data that measure academic growth and inform instruction.

According to NWEA (2012c), computer-adaptive tests are designed to adjust to the student's performance level. The level of difficulty for each question is based on the student's ability level and performance on the previous question. If a student answers a question incorrectly, the level of difficulty decreases; if a student answers a question correctly, the level of difficulty increases. Schools use computer-adaptive assessments because they provide detailed diagnostic information about student performance (Shapiro & Gebhardt, 2012). According to NWEA (2012c) there are many advantages to administering computer-adaptive tests: (a) test scores are available immediately, (b) tests are adaptive and thus paced based on the individual student, (c) test security is increased because no hard copy is available, (d) test time is reduced because fewer items are needed to provide precise scores, and (e) accurate test scores are available for a wide range of ability levels. The adaptive nature of NWEA assessments allows educators to determine

the instructional level of individual students and measure academic growth over time regardless of age or grade level.

NWEA-MAP assessments are computer-adaptive assessments created by NWEA designed to provide detailed results that allow teachers to individualize instruction based on a student's educational needs. MAP assessments are adaptive; therefore, they adapt to students' responses. For example, if a student answers a question correctly, the computer-adaptive assessment provides a more challenging item; however, if a student answers a question incorrectly, it provides a less challenging item.

MAP assessments provide teachers with detailed information about what students have learned and are ready to learn next. Currently, NWEA offers reading, mathematics, science and language MAP assessments, and MAP for Primary Grades in reading and mathematics. MAP assessments are aligned to state and national standards (NWEA, 2012d). MAP assessments provide student scale scores based on the Rasch model. Based on student responses to questions on the MAP assessment, students are assigned a specific Rasch Unit or RIT scale score. According to NWEA (2012f), the RIT scale provides an estimate of student achievement. The RIT scale is an accurate achievement scale that has equal intervals despite the grade or age of the individual being assessed and it has been proven to help measure growth over time (NWEA, 2012f).

According to NWEA (2012h), their researchers have collected an extensive amount of evidence over the years to support the reliability of NWEA assessments. Based on their analyses, "test and re-test studies have consistently yielded statistically valid correlations between multiple test events for the same students" (NWEA, 2012h, p. 1). Pearson's correlation coefficient is a statistical term that represents the degree to which

two quantitative values are related (Fraenkel et al., 2012). Pearson's correlation coefficients are values between .00 and 1.00, with values near .00 indicating no linear relationship and values near or at 1.00 indicating a relationship (Fraenkel et al., 2012). Based on the mathematics ($r = 0.78$) and reading ($r = 0.74$) Pearson's correlation coefficients, the MAP reading and mathematics assessments and Mississippi state test for language arts and mathematics are highly related.

Moreover, NWEA researchers have established status and growth norms for each of their MAP assessments. Using the RIT scores from 5.1 million students from across the United States, NWEA data specialists have calculated the mean RIT score growth and status norm of students assessed in Language Usage, Mathematics, General Science, Science Concepts and Processes, and Reading Measures of Academic Progress (NWEA, 2014). According to NWEA (2014), extensive research methods were used to ensure that both status and growth norms were representative of the American school-age population. Educators are able to use a student's status norm to identify the student's percentile rank for any instructional week of the school term; this allows an educator to compare a student's performance to the performance of a national sample of students in a particular subject area (NWEA, 2014). Additionally, educators may use the growth norms to establish growth targets for students (NWEA, 2014). Tables 7 and 8 represent the reading and mathematics norms established during the 2011 norms study.

Table 7

2011 Reading Status Norms

Grade	Beginning of Year	Middle of Year	End of Year
K	142.5	150.6	156.0
1	160.3	170.7	176.9
2	175.9	183.6	189.6
3	189.9	194.6	199.2
4	199.8	203.2	206.7
5	207.1	209.8	212.3
6	212.3	214.3	216.4
7	216.3	218.2	219.7
8	219.3	221.2	222.4
9	221.4	221.9	222.9
10	223.2	223.4	223.8
11	223.4	223.5	223.7

Table 8

2011 Mathematics Status Norms

Grade	Beginning of Year	Middle of Year	End of Year
K	143.7	150.5	156.1
1	162.8	172.4	179.0
2	178.2	185.5	191.3
3	192.1	198.5	203.1
4	203.8	208.7	212.5
5	212.9	217.8	221.0
6	219.6	222.8	225.6
7	225.6	228.2	230.5
8	230.2	232.8	234.5
9	233.8	234.9	236.0
10	234.2	235.5	236.6
11	236.0	237.2	238.3

NWEA provides two types of information to educators so that they may link NWEA assessment results to state curriculum content: Scale Alignment Studies and Goal Structures (NWEA, 2012b). Scale Alignment Studies link the RIT scales to proficiency levels for each subject (i.e., reading, mathematics, science) and grade level of state

assessments (NWEA, 2012b). Goal Structures link curriculum content from state standards to reporting areas of a state aligned MAP assessment (NWEA, 2012b).

Scale Alignment Studies, also called Linking Studies, investigate the connection between MAP assessments and state standardized achievement assessments (NWEA, 2012g). NWEA researchers study the performance of students who have completed both MAP and state standardized achievement assessments to determine the relationship between the two assessments (NWEA, 2012g). "Each study identifies the specific RIT scale scores from MAP that correspond to the various proficiency levels for each subject and student grade" (NWEA, 2012g, p. 1). These studies provide estimates of the probability that a student who obtains a specific RIT score will achieve a status of proficient or greater on his/her state standardized achievement assessment (NWEA, 2012g). State standardized assessments vary from state-to-state; therefore, different linking studies are necessary for each state (NWEA, 2012g). Moreover, NWEA uses the Equipercentile Method to estimate MAP RIT cut scores that correspond to specific proficiency levels on state achievement assessments (NWEA, 2012g).

NWEA alignment studies describe the connection between MAP assessments and state standardized achievement assessments. The most recent Mississippi Scale Alignment Study was completed in February of 2011. The results of this study indicated a strong correlation between MAP RIT scale scores and the MCT2 (NWEA, 2011). In their analysis, NWEA used the MCT2 results from a sample of 22,483 students in grades 2 – 8 from 64 Mississippi schools who took the state assessment and the NWEA-MAP assessment during the 2009-2010 school term (NWEA, 2011). Using the Equipercentile

Method, minimum NWEA RIT cut scores that correspond to specific MCT2 proficiency levels were established (NWEA, 2012g).

Tables 9 and 10 demonstrate "the best estimate of the minimum RIT equivalent to each state performance level" (i.e., basic, proficient, advanced) for Grades 3 – 8 based on the students' spring mathematics and reading RIT scores (NWEA, 2011, p. 2). Students who achieve at least the minimum mathematics or reading RIT cut score for each performance level category have a 50% probability of achieving the corresponding performance level on the MCT2 (NWEA, 2011). The information presented in these tables can be used to identify students who demonstrate academic difficulties.

Table 9

Minimum Estimated Spring RIT Cut Scores Corresponding to MCT2 Performance Levels – Mathematics

	Cut Scores and Percentiles for each State Performance Level						
	Below Basic	Basic		Proficient		Advanced	
Grade	Cut Score	Cut Score	Percentile	Cut Score	Percentile	Cut Score	Percentile
2	<175	175	9	186	35	201	82
3	<185	185	9	198	35	213	82
4	<196	196	13	206	35	227	87
5	<203	203	14	213	34	236	87
6	<206	206	14	219	37	240	86
7	<210	210	15	222	34	245	84
8	<212	212	13	224	29	247	79

Note. The cut scores shown in the table are the minimum estimated scores. Meeting the minimum MAP cut score corresponds to a 50% probability of achieving that performance level. Use the probabilities in Tables B1 – B4 to determine the appropriate 'target' scores for a desired level of certainty.

Table 10

Minimum Estimated Spring RIT Cut Scores Corresponding to MCT2 Performance Levels – Reading

Cut Scores and Percentiles for each State Performance Level							
	Below	Basic		Proficient		Advanced	
Grade	Cut Score	Cut Score	Percentile	Cut Score	Percentile	Cut Score	Percentile
2	<174	174	13	189	48	203	84
3	<183	183	13	199	48	212	84
4	<188	188	11	204	42	219	85
5	<195	195	12	210	43	227	90
6	<197	197	10	213	41	232	91
7	<199	199	10	217	44	237	94
8	<207	207	15	222	49	240	94

Note. The cut scores shown in the table are the minimum estimated scores. Meeting the minimum MAP cut score corresponds to a 50% probability of achieving that performance level. Use the probabilities in Tables B1 – B4 to determine the appropriate 'target' scores for a desired level of certainty.

Presented in Appendix B, Tables B1-B4, NWEA researchers also calculated the estimated probabilities of a student achieving proficiency on the state standardized assessment, based on his/her observed RIT Score (NWEA, 2011). This information can be used to categorize RIT specific objectives that seemingly correlate to a student's level of performance on the state test; moreover, these results can also be used to identify students who are at-risk of not passing the state standardized assessment (NWEA, 2011).

According to the results of the 2011 Mississippi Scale Alignment Study, NWEA researchers were able to accurately predict the mathematics state test performance of eighth grade students 82.5% of the time using their observed MAP score; the researchers were able to accurately predict the language arts state test performance of eighth grade students 80.4% of the time using their observed MAP score (NWEA, 2011). Located in Appendix B, Table B5 represents the percentage of eighth grade students NWEA

researchers accurately predicted, underestimated and overestimated state standardized test performance (NWEA, 2011).

Many schools are choosing to use NWEA-MAP as a viable assessment option (Merino & Beckman, 2010). NWEA-MAP assessments allow schools to use an adaptive assessment that provides academic information that may be used in a timely manner. After a student has taken a MAP assessment, teachers are able to review instructional data that outlines what a student knows and is ready to learn next. NWEA-MAP assessments provide each student with a RIT score that corresponds to a specific MCT2 proficiency level; schools and districts may use this information to identify students who are in need of remediation.

Mississippi Statewide Accountability System and Mississippi Curriculum Test

Mississippi developed its first assessment systems as early as 1982. Since then, Mississippi has improved and strengthened its curriculum frameworks and accountability systems to meet the requirements of state and federal legislations. The current accountability system is the fourth to be utilized in Mississippi. Current assessments that comprise the Mississippi Statewide Accountability System are Mississippi Alternate Assessment of Extended Curriculum Frameworks (MAAECF), Mississippi Science Test (MST2), SATP2, MCT2, Mississippi Writing Assessment Program, and Subject Area Alternative Assessment (SATP2 AA).

A few important facets of the Mississippi Statewide Accountability System include the Achievement, Growth, AYP Models and High School Completion Index (MDE, 2010). Schools, districts, and states are held accountable for student performance under Title I of NCLB based on AYP (U.S. Department of Education, 2002a). According

to MDE (2010), the Achievement Model is defined as a measure of the previous school term's overall school or district level performance. The Growth Model is a measure of the degree to which a school or district meets its expected performance during the previous school term (MDE, 2010). Schools and districts are assigned a performance classification based on the results from the Achievement and Growth Models. According to MDE (2012b), performance classifications are: (a) A – Star School, (b) B – High Performing, (c) C – Successful, (d) D – Academic Watch, (e) F – Low Performing, (f) F – At-Risk of Failing, and (g) F – Failing.

Using standardized achievement assessment data, the Achievement Model yields the Quality of Distribution Index (MDE, 2010). MDE uses the Quality of Distribution Index to measure school and district achievement (MDE, 2010). The Quality of Distribution Index formula is QDI = % at Basic + (2 x % at Proficient) + (3 x % at Advanced) (MDE, 2012d). The minimum Quality of Distribution Index is zero and the maximum is 300. Mississippi's Quality Distribution Index for the 2011-2012 school term is 162, which falls within the Successful performance classification.

The Growth Model provides a Growth Composite that predicts students' standardized achievement assessment performance for the following year (MDE, 2010). The Growth Composite is computed using students' assessment data from the previous two school terms (MDE, 2010). The High School Completion Index is an important accountability factor for high schools and districts with a graduating 12th grade high school class. High school completion is calculated by tracking a cohort of students beginning their ninth grade year and continuing to track those students for up to five

years (MDE, 2010). High school completion and graduation rate are used to calculate the High School Completion Index (MDE, 2010).

A school or district receives its annual Accountability Status or Performance Classification based on the Achievement Model, Growth Model, and, if applicable, High School Completion Index (MDE, 2010). Title I of NCLB requires schools and districts to meet AYP requirements, but Mississippi has requested flexibility under revised legislation proposed by President Obama (MDE, 2012c).

According to MDE (2010), NCLB requires school, district, and state report cards. Those report cards are required to contain specific information that falls into three categories: school improvement, teacher qualifications, and test data. Report cards must contain specific information about the professional qualifications of core academic subject area teachers. Student achievement data must be disaggregated by race, ethnicity, gender, disability status, migrant status, English proficiency, and economically disadvantaged status. Report cards should list a comparison of each subgroup's achievement levels. States are required to report the percentage of students not tested and the most recent 2-year student achievement data. Graduation rates must also be disaggregated in the same way as achievement test data.

MDE uses the MAARS database to provide state, district, and school report card data (MDE, 2010). The state report card provides a detailed overview of all student assessment data in Mississippi (MDE, 2013b). The state report card includes student enrollment by grade level, student enrollment by gender and ethnicity, and state poverty level. The state report card lists the number of students tested, mean scale score, and percentage of students scoring at each performance level (minimal, basic, proficient,

advanced) by grade level. The state report card includes, by grade level, the percentage of students scoring basic or above and proficient or above by grade level for each standardized assessment (e.g., MCT2, MST2, Writing, SATP2) administered in the state. Additionally, under the new accountability system where Mississippi has requested flexibility from NCLB, Mississippi reports the details of AMO goals by subject area and subgroup (limited English proficient, students with disabilities, economically disadvantaged, ethnicity); the district and individual schools' Differentiated Accountability Label (i.e., approaching target, focus, high performing reward, high progress reward, on target, priority) are also reported. Moreover, the state report card lists the QDI of the state, and each Mississippi district and schools within each district.

The district report card provides a detailed overview of all student assessment data within a Mississippi school district. The district report card includes, by district, student enrollment by grade level, student enrollment by gender and ethnicity, and district poverty level. The district report card lists the number of students tested, mean scale score, and percentage of students scoring at each performance level (minimal, basic, proficient, advanced) by grade level. The district report card includes, by grade level, the percentage of students scoring basic or above and proficient or above by grade level for each standardized assessment (e.g., MCT2, MST2, Writing, SATP2, SATP2-AA) administered in the district. The district level report card lists the details of Annual Measurable Objectives by subject area and subgroup (limited English proficient, students with disabilities, economically disadvantaged, ethnicity), as well as, the district and individual schools' Differentiated Accountability Label (i.e., approaching target, focus, high performing reward, high progress reward, on target, priority). In addition, district

report cards provide school district accreditation status and district Achievement and Growth Model data: accountability status, AMO, growth status, graduation rate, and High School Completion Index. In addition, the district report card lists the QDI of each school within the specific district.

The school report card provides a detailed overview of all student assessment data within a specific Mississippi school. The school report card includes, by school, student enrollment by grade level, student enrollment by gender and ethnicity, and school poverty level. The school report card lists the number of students tested, mean scale score, and percentage of students scoring at each performance level (minimal, basic, proficient, advanced) by grade level. The school report card includes, by grade level, the percentage of students scoring basic or above and proficient or above by grade level for each standardized assessment (e.g., MCT2, MST2, Writing, SATP2, SATP2-AA) administered in the school. The school report card lists school level AMO data as required by Title I of NCLB; AMO reporting requirements include reporting whether AMO goals were met by subject area and subgroup (i.e., limited English proficient, students with disabilities, economically disadvantaged, ethnicity). In addition, school report cards provide school level Achievement and Growth Model data: accountability status, QDI, growth status, graduation rate, and High School Completion Index; graduation rate and High School Completion Index are provided for schools with graduating twelfth grade students.

The Mississippi Statewide Accountability System includes MAAECF, MST2, SATP2, MCT2, Mississippi Writing Assessment Program, and SATP2-AA. Schools and districts are assigned performance levels based largely on student assessment scores. The

focus of this study will be the MCT2; therefore, the researcher will present detailed information about this assessment only in the next section.

Mississippi Curriculum Test, Second Edition (MCT2)

The MCT2 is a criterion-referenced assessment consisting of tests of language arts and mathematics. Items on the language arts MCT2 are aligned with the 2006 Mississippi Language Arts Framework – Revised and items on the mathematics MCT 2 are aligned with the 2007 Mississippi Mathematics Framework – Revised. Students in grades 3 through 8 are administered the MCT2 each year in May. Based on student results, individual student scale scores and proficiency levels are assigned. The proficiency levels were selected based on cut scores identified by committees of Mississippi teachers and approved by MDE. There are four proficiency levels: Advanced, Proficient, Basic, and Minimal (MDE, 2010).

According to the MCT2 Interpretive Guide, the MCT2 measures achievement in Grades 3 – 8 in language arts and mathematics (MDE, 2011). The test items presented on the MCT2 range in degree of difficulty based on Mississippi's academic standards. The language arts and mathematics frameworks are organized into competencies for each grade level and subject. Competencies are learning standards that are required at each grade level. For each competency, there are objectives. The objectives list the skills that are required to attain competencies, explain the competencies thoroughly, or show the evolution of content for each grade level (MDE, 2011).

The MCT2 language arts assessment meets the federal guidelines of NCLB and is a measure of achievement for grades 3 – 8 based on the 2006 Mississippi Language Arts

Curriculum Framework – Revised (MDE, 2011). According to MDE (2011), the following competencies are used to measure language arts achievement:

1. *Vocabulary*: The student will use word recognition and vocabulary (word meaning) skills to communicate.
2. *Reading*: The student will apply strategies and skills to comprehend, respond to, interpret, or evaluate a variety of texts of increasing length, difficulty, and complexity.
3. *Writing*: The student will express, communicate, evaluate, or exchange ideas effectively.
4. *Grammar*: The student will apply Standard English to communicate.

The MCT2 mathematics assessment meets the federal guidelines of NCLB and is a measure of achievement for grades 3 – 8 based on the 2007 Mississippi Mathematics Curriculum Framework – Revised (MDE, 2011). The MCT2 mathematics assessment for grades 3 – 7 is a measure of general mathematics achievement; MCT2 mathematics for grade 8 is a measure of Pre-Algebra achievement. According to MDE (2011), the following competencies are used to measure mathematics achievement:

1. *Number and Operations*: Analyze relationships among numbers and the four basic operations. Compute fluently and make reasonable estimates.
2. *Algebra*: Explain, analyze, and generate patterns, relationships, and functions using algebraic symbols, demonstrate an understanding of the properties of the basic operations, and analyze change in various contexts.
3. *Geometry*: Develop mathematical arguments about geometric relationships and describe spatial relationships using coordinate geometry.

4. *Measurement*: Develop concepts and apply appropriate tools and techniques to determine units of measure.
5. *Data Analysis and Probability*: Formulate questions that can be addressed with data and select and use appropriate statistical methods to analyze data. Apply basic concepts of probability.

The MCT2 performance levels are ordered into four levels. The performance levels describe the content and processes at each level that a student is expected to know, demonstrate, or perform (MDE, 2011). The performance levels are:

1. *Minimal*: Students at the minimal level inconsistently demonstrate the knowledge or skills that define basic level performance. These students require additional instruction and remediation in the knowledge and skills that are necessary for success in the grade or course in the content area.
2. *Basic*: Students at the basic level demonstrate partial mastery of the knowledge and skills in the course and may experience difficulty in the next grade or course in the content area. These students are able to perform some of the content standards at a low level of difficulty, complexity, or fluency as specified by the grade-level content standards. Remediation is recommended for these students.
3. *Proficient:* Students at the proficient level demonstrate solid academic performance and mastery of the knowledge and skills required for success in the grade or course in the content area. These students are able to perform at the level of difficulty, complexity, or fluency specified by the grade-level content standards. Students who perform at this level are

prepared to begin work on even more challenging material that is required in the next grade or course in the content area.

4. *Advanced*: Students at the advanced level consistently perform in a manner clearly beyond that required to be successful in the grade or course in the content area. These students are able to perform at a high level of difficulty, complexity, or fluency as specified by the grade-level content standards (MDE, 2012a).

Based on individual student performance, students who take the MCT2 are assigned a scale score; each scale score corresponds to a proficiency level. The scale score ranges and corresponding proficiency levels as described in the MCT2 Interpretive Guide for each MCT2 performance level in Grades 3 – 8 are described in Tables 11 and 12 (MDE, 2011).

Table 11

Mississippi Student Performance Standards Performance Levels Scale Score Ranges for MCT2 Language Arts

	Label			
Grade	Advanced	Proficient	Basic	Minimal
3	162 and above	150-161	138-149	137 and below
4	162 and above	150-161	138-149	137 and below
5	164 and above	150-163	138-149	137 and below
6	166 and above	150-165	137-149	136 and below
7	168 and above	150-167	138-149	137 and below
8	167 and above	150-166	138-149	137 and below

Note. The lowest and highest possible attainable scale scores will vary as new forms of the assessment are developed.

Table 12

Mississippi Student Performance Standards Performance Levels Scale Score Ranges for Mathematics

	Label			
Grade	Advanced	Proficient	Basic	Minimal
3	165 and above	150-164	138-149	137 and below
4	165 and above	150-164	141-149	140 and below
5	164 and above	150-163	141-149	140 and below
6	164 and above	150-163	142-149	141 and below
7	164 and above	150-163	142-149	141 and below
8	164 and above	150-163	142-149	141 and below

Note. The lowest and highest possible attainable scale scores will vary as new forms of the assessment are developed.

The MCT2 is an important facet of the Mississippi Statewide Accountability System. All students in Grades 3 – 8 are administered the MCT2. Based on their performance students are assigned a scale score that corresponds to a specific performance level. Student scale scores are used in the calculation to assign schools and districts to performance levels. School and district performance levels hold many

implications. Schools and districts that obtain favorable performance levels may be entitled to additional funding, awards, and public praise. Schools and districts that obtain unfavorable performance levels may be closed, taken over by charter networks, or subject to public condemnation.

Summary

NCLB requires that students meet state-standards in reading and mathematics and are proficient by 2014 (U.S. Department of Education, 2002a). States are responsible for developing assessment and accountability systems that specify how each state plans to make sure all students receive a fair and equal opportunity to obtain a high quality education. An integral part of each state's assessment and accountability system is their state mandated standardized achievement assessments. States use their state standardized achievement assessments to measure student proficiency.

Mississippi's third through eighth grade students are required to take the MCT2 each year in May. Assessment results yield significant implications for students, teachers, schools and districts. As students obtain favorable scores on the MCT2, those implications are positive with students, teachers, schools, and district earning rewards, receiving positive praise, and often schools and districts may receive additional funding. On the contrary, as students obtain unfavorable scores on the MCT2, negative consequences may be imposed for students, teachers, and schools; those negative consequences may include restricted curriculum offerings for students; teachers may face termination or re-assignment, while schools and districts may receive sanctions, loss of funding, or even closure.

Because of the significant implications of state standardized assessment results, states are searching for ways to ensure students are prepared for state standardized assessments. NWEA-MAP assessments are designed to be administered three times a year, fall, winter, and spring. NWEA asserts that their MAP assessments will provide students, teachers, schools, and districts with the prescriptive information needed to ensure students are prepared for state mandated achievement assessments. NWEA also asserts that their MAP assessment may be used to predict student success on the MCT2.

In this study, the researcher used multiple regression analysis to determine the predictor value of the NWEA-MAP assessment on the MCT2. Specifically, if the reading and mathematics NWEA-MAP assessments can accurately predict students' language arts and mathematics MCT2 scale scores, Mississippi students, teachers, schools, and districts may have a complementary assessment that is able to provide them with (a) immediate test scores that are aligned with the MCT2, (b) results from an adaptive assessment that gives prescriptive information based on individual student results, (c) results from an assessment that ensures increased security because no hard copy of the assessment is available, (d) results in a timely manner because test time is reduced because fewer items are needed to provide precise scores, (e) accurate test scores that are available for a wide range of ability levels, and (f) the option to obtain prescriptive instructional information based on assessment data that is aligned to the state standards up to three times per year (fall, winter, spring) versus only at the end of the school term when the MCT2 data are available.

CHAPTER III

METHODS

Introduction

The purpose of this study was to determine whether reading and mathematics NWEA-MAP assessments are an adequate measure to predict student achievement on the MCT2. Although NWEA-MAP assessments are not required for state accountability purposes, many school districts administer the assessments to identify students who are at-risk of not performing well on their end-of-year state mandated assessment, as well as, for the prescriptive and diagnostic instructional information provided for each student who takes the assessments.

Specifically, the researcher sought to determine whether sixth and eighth grade students' performance on the language arts MCT2 could be predicted using the students' fall and spring reading NWEA-MAP RIT scores. The researcher also sought to determine whether sixth and eighth grade students' performance on the mathematics MCT2 could be predicted using the students' fall and spring mathematics NWEA-MAP RIT. The researcher also examined whether student characteristics of gender, ethnicity, and economically disadvantaged status predicted students language arts and mathematics MCT2 performance.

The following research questions guide this study:

1. How accurately do fall NWEA-MAP RIT reading scores predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade?
2. How accurately do fall NWEA-MAP RIT mathematics scores predict mathematics achievement scores as measured by the MCT2 mathematics scale scores for students in sixth and eighth grade?
3. How accurately do spring NWEA-MAP RIT reading scores predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade?
4. How accurately do spring NWEA-MAP RIT mathematics scores predict mathematics achievement scores as measured by the MCT2 mathematics scale scores for students in sixth and eighth grade?
5. How accurately do the fall and spring NWEA-MAP RIT reading scores, gender, ethnicity and economically disadvantaged status predict student performance on the MCT2 language arts assessment for students in sixth and eighth grade?
6. How accurately do the fall and spring NWEA-MAP RIT mathematics scores, gender, ethnicity and economically disadvantaged status predict student performance on the MCT2 mathematics assessment for students in sixth and eighth grade?

This chapter discusses the participants and setting, instrumentation, research design, procedures, and methods of data analysis.

Participants and Setting

Population

It was the researcher's intention to study the data from a school district where student enrollment is similar to the demographic make-up of the state of Mississippi. Table 13 demonstrates the similarities in the demographic make-up for Mississippi and the school district used for this study.

Table 13

Demographic Comparison of Mississippi and School District

Subgroup	Mississippi Percentage	School District Percentage
Asian	0.9	3.2
African American	37.0	30.4
Caucasian	59.1	62.9
Hispanic	2.9	3.0
Native American	0.5	0.2
Pacific Islander	0.0	0.1

Note. Mississippi demographic information was taken from the United States Census Bureau website. School district demographic information was taken from the Mississippi Department of Education website.

There were 13,833 students enrolled in the district during the 2011-2012 school term. Table 14 presents the demographic information by subgroup for this district. For the purposes of this study, the researcher will use three subgroups for the ethnicity category, African American, Caucasian, and Other.

Table 14

2011 – 2012 District Enrollment by Subgroup

Subgroup	Enrollment	Percentage
Female	6717	48.5
Male	7116	51.5
African American	4213	30.4
Caucasian	8696	62.9
Other	924	6.7
Economically Disadvantaged	804	5.8

Sample

The sample used for this study consists of students in grades 6 and 8 who attend a Southern Mississippi school district. The district has three elementary and/or middle schools that contain sixth and eighth grade students and one school that is an eighth grade only school. According to the District Report Card published by the Mississippi Department of Education, during the 2011-2012 school term, 953 eighth grade students from the school district took the language arts and mathematics MCT2 in May of 2012; there were 1089 sixth grade students from the school district who took the language arts assessment and 1085 sixth grade students who took the mathematics MCT2 assessment in May of 2012 (MDE, 2013b).

The participants in this study include only sixth and eighth grade students who took the NWEA-MAP reading and mathematics assessments that were administered during the fall and spring testing cycles and the language arts and mathematics MCT2 in May during the 2011-2012 school term. Data for 676 sixth grade students were used in this study. Data for 659 eighth grade students were used in this study.

Sample size. Mertler and Vannatta (2010) suggest that the ratio of participants to independent variables be at least 15 to 1; this will "provide a reliable regression equation"

(p. 163). According to Fraenkel et al. (2012), correlational studies should have a sample size of at least 30 in order to provide results that are of substance. A study of this nature, with five independent variables (Reading NWEA-MAP RIT, Mathematics NWEA-MAP RIT, Gender, Ethnicity, Economically Disadvantaged status), should have a sample size of at least 75 to provide a reliable regression equation.

Within the district, there are four schools from which student NWEA-MAP and MCT2 data are used. Only the data from students who took the reading and mathematics NWEA-MAP assessment during the fall and spring and the language arts and mathematics MCT2 during the 2011-2012 school term are used in the analysis. Data from 676 sixth grade students and 659 eighth grade students make up the representative sample for this study.

Instrumentation

Two instruments, the MCT2 and the NWEA-MAP, were used to gather data to answer the six research questions that guide this study. The dependent variables used in this study were sixth and eighth grade students' MCT2 language arts and mathematics scale scores. The fall and spring language arts and mathematics MAP assessment data, along with student characteristics of gender, ethnicity, and economically disadvantage status were used as the independent or predictor variables for this study.

The MCT2 was designed to meet the testing requirements mandated by the NCLB Act. The MCT2 consists of criterion-referenced language arts and mathematics assessments. NCLB requires all students in Grades 3 – 8 to take the MCT2 in May of each school term. These assessments are aligned with the 2006 Mississippi Language Arts Framework-Revised and the 2007 Mississippi Mathematics Framework-Revised.

Students in Grades 3 – 8 are administered the MCT2; special education students in Grades 3 – 8 whose Individualized Education Plan (IEP) specify instructional goals that are aligned with the 2006 Mississippi Language Arts Framework-Revised and the 2007 Mississippi Mathematics Framework-Revised are also administered the MCT2. The results are utilized to improve instruction and accelerate student achievement. In addition, the results are used in the Achievement Model, Growth Model, and Annual Measurable Objectives (MDE, 2012b).

According to the most recent MCT2 Technical Manual update (Pearson, 2011), there are a total of 83 test items, per test form, with 70 core items and 13 non-core field test items on the eighth grade language arts MCT2 assessment; on the eighth grade mathematics MCT2 assessment there are a total of 60 test items per test form, with 50 core items and 10 non-core items. Test scores are based on the core test items only. The sixth grade mathematics assessment consists of 60 test items, per test form, with 50 core items and 10 non-core field test items; the sixth grade language arts assessment consists of 73 test items, per test form, with 60 core items and 13 non-core field test items (Pearson, 2011).

The MCT2 is an untimed, multiple choice assessment that requires all students to bubble in answers on the provided answer document. The MCT2 consists of two tests: language arts and mathematics. There are multiple forms of each test. The language arts test has two parts: reading and writing. The reading and writing portions of the language arts test are administered on separate days, with the language arts test lasting for two days. The mathematics test is also administered over the course of two days. On the mathematics test, eighth grade students are allowed to use a calculator, 12-inch lead-in

edge ruler with English and Metric Measurements, and are provided a formula chart to use on the Pre-Algebra test.

MCT2 results are reported as scale scores and each scale score corresponds to a performance level. The current MCT2 performance levels are Advanced, Proficient, Basic, and Minimal. For the purpose of this study, the researcher will determine whether NWEA-MAP reading and mathematics RIT scores can predict MCT2 language arts and mathematics scale scores.

The MCT2 test developer, Pearson, coordinates test administration procedures with the Mississippi Department of Education. Pearson provides District Test Coordinator Manuals, School Test Coordinator Manuals, and Test Administrator Manuals for each individual responsible for test administration at the appropriate level before test administration; each manual is to be returned on or before a specified deadline after testing is completed (Pearson, 2008).

Using the train-the-trainer model, Pearson provided the District Test Coordinators with test administration training and the District Test Coordinators are expected to provide all other individuals who will coordinate or administer the MCT2 assessment with the appropriate training. Training includes a standardized procedure to inventory, handle, administer and package MCT2 materials. District Test Coordinators are responsible for coordinating and monitoring testing procedures for each school within their school district (Pearson, 2008).

School Test Coordinators are responsible for monitoring all aspects of test administration within their respective buildings. They are responsible for following protocols described in the School Test Coordinator's Manual in regards to how to

administer, inventory, handle, package test materials and respond to situations that may arise during testing. To ensure standardization, test administrators follow the instructions presented in the Test Administrator's Manual and read aloud all directions and information to students as specified by the manual (Pearson, 2008).

According to Pearson (2008), students should be tested in a comfortable environment, with suitable lighting, sufficient work space, and a quiet setting; it is very important for students to be tested under ideal physical conditions. Test Administrators are required to remove or cover all content related materials form the testing environment. Provisions should be made to ensure that students with Individual Education Plans or Section 504 Individual Accommodation Plans receive allowable testing accommodations. Other provisions that require advanced preparation include ensuring that arrangements have been made for students who do not finish testing by the end of the test administration time and providing a monitor for students who suspend testing in order to eat lunch to ensure students do not discuss MCT2 items (Pearson, 2008).

According to Pearson (2008), in accordance with the NCLB policy, students identified as having a significant medical emergency and are unable to participate in MCT2 testing could be exempt from participation rate calculations. Those students must meet three criteria:

> The student's non-participation in the regularly scheduled testing or makeup testing was due to circumstances beyond the control of the school district, the student was determined by a medical practitioner to be so incapacitated as to be

unable to participate in the state assessment, and the medical emergency was due to unforeseen events or situations (Pearson, 2008, p. 21).

District and school test coordinators are responsible for the secure collection, packaging and return of MCT2 test materials for scoring. MCT2 results are made available to the state, district and school after Pearson has scored student assessments. The results are presented electronically to the state, district, and school. Students and parents receive paper copies of individual student results.

Reliability and validity information pertaining to the MCT2 is published in the MCT2 Technical Manual. According to Pearson (2008), MCT2 items were analyzed for construct, criterion, concurrent, and content validity. Reliability was also measured using Cronbach's alpha for all students and selected subgroups. Table 15 describes reliability statistics for Grades 6 and 8. It was concluded that the language arts and mathematics MCT2 are both reliable and valid measures.

Table 15

Reliability (Cronbach's alpha) of MCT2 Tests

Subject	Grade	N	All	Female	Male	African American	Caucasian	Economically Disadvantaged	
								No	Yes
Language Arts	6	59	0.84	0.83	0.84	0.81	0.84	0.84	0.77
Language Arts	8	68	0.87	0.87	0.87	0.84	0.87	0.87	0.75
Mathematics	6	49	0.91	0.90	0.91	0.88	0.90	0.90	0.85
Mathematics	8	50	0.90	0.89	0.90	0.87	0.90	0.90	0.80

According to NWEA, the MAP assessment is a computerized adaptive assessment designed to assist with classroom instruction by presenting teachers with the knowledge of what students know and what students are ready to learn. Each test item on a MAP assessment corresponds to a value on the RIT scale. MAP offers assessments in mathematics, reading, language, and science. MAP assessments provide detailed data about what a student has learned academically and what a student needs to learn to move forward academically or obtain a higher RIT score. NWEA provides on-site and online training for MAP administration. Training is provided to district and school level MAP coordinators, proctors, teachers and others who will administer the assessment or use the assessment data. Districts may select training topics based on the version of MAP they use: web-based or client server. Topics include:

1. *MAP Basics*: MAP system and NWEA assessments foundation.
2. *MAP Proctor Training*: Skills, knowledge, and resources necessary for successful proctoring.

3. *Enrolling for a Test Term*: How to create and submit MAP testing rosters and upload to the NWEA website. Rosters must be submitted each season of testing (fall, winter, spring).
4. *Using Network Test Environment (NTE) Administration*: Client Server MAP Administration requires the use of a NTE folder. The NTE folder is a folder that stores and hosts student and test database information. NTE Administration software is a Windows-based application used to perform the download and upload of tests, test results and student data to and from the NTE folder. The NTE folder must exist on a shard network accessible to local user accounts on the testing workstations.

MAP Coordinators are responsible for uploading MAP testing rosters each season of testing. MAP Proctors are responsible for setting up student computers for MAP assessments. When using the client server version, Proctors download the TestTaker application to each computer that will be used for MAP testing. TestTaker is an application that retrieves questions from the NTE and sends students' responses back to the NTE for recording. Once the TestTaker application is downloaded to each computer, the Proctor prepares the workstation by selecting the correct test and student's name. It is recommended that test results be uploaded through the NTE Administration application daily. Student results, in the form of RIT scores, are typically available within 24 hours of the NTE Administration upload.

According to NWEA (2012h), their researchers have collected an extensive amount of evidence over the years to support the reliability and validity of NWEA assessments. NWEA researchers have analyzed the results of thousands of students in

several states to determine if their assessments are reliable and possess content, concurrent, predictive, and criterion-related validity. NWEA determined that their MAP assessments are both reliable and valid. According the NWEA-MAP Technical Manual, NWEA-MAP assessments have also been analyzed to determine the correlation to state content aligned assessments such as the MCT2. Using Pearson's correlation coefficient, NWEA researchers have concluded that NWEA-MAP assessments are highly correlated to MCT2 assessments. Table 16 demonstrates the reading and mathematics correlations for sixth and eighth grade students.

Table 16

Pearson's correlations for State Content Aligned MAP Reading and Mathematics Assessments

Test	Grade 6	Grade 8
MAP Reading Survey with Goals 6+	0.753	0.751
MAP Mathematics Survey with Goals 6+	0.870	0.822

Research Design

This study used a correlational research design. Correlational research may be used to predict outcomes (Fraenkel et al., 2012). Correlational research is also used to describe relationships among variables (Fraenkel et al., 2012). The authors note that when variables correlate, "scores within a certain range on one variable are associated with scores within a certain range on the other variable" (Fraenkel et al., 2012, p. 331).

According to Fraenkel et al. (2012), positive correlations are found when high scores on one variable are related to high scores on the other variable and similarly low scores on one variable are related to low scores on the other variable; negative correlations are found when high scores on one variable are related to low scores on the

other variable and low scores on one variable are related to high scores on the other variable.

If a strong relationship is found to exist between variables, it is possible for the researcher to "predict a score on one variable if a score on the other variable is known" (Fraenkel et al., 2012, p. 333). The predictor variable is used to make the prediction, while the variable for which the prediction is made is the criterion variable (Fraenkel et al., 2012). Multiple regression is a more complex correlational method that can be used to determine a correlation between a criterion variable and the best combination of two or more predictor variables (Fraenkel et al., 2012).

The independent variables are NWEA-MAP reading and mathematics RIT scores and student characteristics of gender, ethnicity, and economically disadvantaged status. For the purposes of this study, each of the independent variables were operationally defined as: (a) NWEA-MAP assessments are optional computer-adaptive assessments that schools are not required to administer by state or federal education mandates; after completing a MAP assessment, students obtain a RIT score that corresponds to a specific MCT2 proficiency level; (b) gender refers to whether a student is male or female; (c) ethnicity refers to the race subgroup each student within a school district is assigned; for the purposes of this study, the researcher will reference the following ethnicity subgroups: African American, Caucasian, and other; and (d) economically disadvantaged status refers to a students' free/reduced or full price lunch status; students who receive free/reduced lunch are considered economically disadvantaged. The dependent variables are language arts and mathematics MCT2 scale scores.

Procedures

The researcher received approval for the current research from her doctoral committee, Southern Mississippi school district, and Mississippi State University Institutional Review Board (See Appendix A). After receiving approval from all parties, the researcher received de-identified individual NWEA-MAP RIT scores and MCT2 scale scores and proficiency levels in an Excel file from the school district. The data included sixth and eighth grade students' subgroup information such as gender, economic disadvantaged status, and ethnicity. A school district issued identification number, that is random and has no personal information that relates to individual students, links each student to individual MCT2 and NWEA-MAP scores. The data were provided to the researcher in an Excel file. Once the data were obtained, the researcher eliminated the scores of all students who did not take all four assessments (NWEA-MAP reading, NWEA-MAP mathematics, MCT2 language arts, MCT2 mathematics). The data were entered into the Statistical Package for the Social Sciences (SPSS) data analysis program.

Data Analysis

Student language arts and mathematics MCT2 scale scores, reading and mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status were summarized using descriptive statistics. Frequencies, means, percentages and standard deviations were reported. Descriptive statistics were used to identify cases with missing data. Cases with missing data were eliminated. Simple linear regression was used to answer the first four research questions and standard multiple regression was used to answer the final two research questions.

Regression Analysis

The primary purpose of regression analysis is to develop an equation that is used to predict values on a dependent variable (Mertler & Vannatta, 2010). Regression analysis can also be used to explain relationships among variables (Mertler & Vannatta, 2010). The researcher used simple linear regression and multiple regression to address the research questions for this study. Simple linear regression is used when there is one dependent variable and one independent variable (Mertler & Vannatta, 2010). Multiple regression is used when there is one dependent variable and multiple independent variables (Fraenkel et al., 2012). Moreover, multiple regression analysis is a statistical method that uses a prediction equation with two or more variables in combination to predict a criterion (Fraenkel et al., 2012).

The researcher assessed the effect of each independent variable on the dependent variable in terms to what it contributed to the prediction of the dependent variable. In order to determine whether fall reading NWEA-MAP RIT scores predict language arts MCT2 scores, a simple linear regression analysis was conducted. In order to determine whether spring reading NWEA-MAP RIT scores predict language arts MCT2 scores, a simple linear regression analysis was conducted. In order to determine whether fall mathematics NWEA-MAP RIT scores predict mathematics MCT2 scale scores, a simple linear regression analysis was conducted. In order to determine whether spring mathematics NWEA-MAP RIT scores predict mathematics MCT2 scale scores, a simple linear regression analysis was conducted. In order to determine whether fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, ethnicity, gender, and economically disadvantaged status predict mathematics MCT2 scale scores,

multiple regression analysis was conducted. In order to determine whether fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, ethnicity, gender, and economically disadvantaged status predict language arts MCT2 scale scores, multiple regression analysis was conducted.

Assumptions of multiple regression. According to Mertler and Vannata (2010), there are two sets of assumptions in multiple regression; the first are assumptions about the raw scale variables:

1. The independent variables are fixed;
2. The independent variables are measured without error;
3. The relationship between the independent variables and the dependent variable is linear (p. 166).

The second set of assumptions is about the residuals (Mertler & Vannatta, 2010):

4. The mean of the residuals for each observation on the dependent variable over many replications is zero;
5. Errors associated with any single observation on the dependent variable are independent of errors associated with any other observation on the dependent variable;
6. The errors are not correlated with the independent variables;
7. The variance of the residuals across all values of the independent variables is constant;
8. The errors are normally distributed (p. 166).

According to Mertler and Vannatta (2010), assumptions 1, 2, and 4 are primarily research design issues, such as sample size and non-metric variables. When considering

sample size, Mertler and Vannatta (2010) suggest that the ratio of participants to independent variables be at least 15 to 1. When using non-metric variables, the researcher must create dummy variables (Nishishiba, Jones, & Kraner, 2013). Assumptions 3, 5, and 6 are concerned with linearity; a linear relationship between the independent variable and the dependent variable indicates that an increase in one variable is associated with a corresponding increase in another variable and a decrease in one variable is associated with a corresponding decrease in another variable (Fraenkel et al., 2012). Assumptions 7 and 8 are concerned with homoscedasticity and normality; homoscedasticity is "the assumption that the variability in scores for one continuous variable is roughly the same at all values of another continuous variable" (Mertler & Vannatta, 2010, p. 33). According to Mertler and Vannata (2010), normality refers to the assumption that all of the observations of a sample are distributed normally.

Multicollinearity can be a problem in correlational research. Multicollinearity happens when moderate to high inter-correlations occur among independent variables that are used in a regression analysis (Mertler & Vannatta, 2010). According to Mertler and Vannatta (2010), a major problem with multicollinearity is that if two independent variables are highly correlated, they contain some of the same information, and are essentially measuring the same thing. Multicollinearity can be addressed by the researcher. The researcher can identify multicollinearity by obtaining tolerance statistics or examining the values for the variance inflation factor (Mertler & Vannatta, 2010). Tolerance is "a measure of collinearity among independent variables where possible values range from 0 to 1"; values close to 0 represent multicollinearity (Mertler & Vannatta, 2010, p. 163). If an independent variable has a tolerance value of less than 0.1,

multicollinearity is an issue (Mertler & Vannatta, 2010). The second method that may be used to assess multicollinearity is to review the values for the variance inflation factor (Mertler & Vannatta, 2010). The values of the variance inflation factor indicate whether there is a strong linear association between the independent variables and the dependent variable; it is suggested that variance inflation factors greater than 10 indicate multicollinearity (Mertler & Vannatta, 2010). In the event that multicollinearity is an issue, the researcher may simply remove the problematic variable from the analysis or "combine the variables involved so as to create a single measure that addresses a single construct, thus deleting the repetition" (Mertler & Vannatta, 2010, p. 163).

When conducting a multiple regression analysis, the researcher should address each of the assumptions. Research design issues such as sample size and dummy coding non-metric variables must also be addressed. According to Mertler and Vannatta (2010), linearity can be assessed by examining bivariate scatterplots; normality can be assessed by examining the values for skewness, kurtosis, and Kolmogorov-Smirnov statistics; homoscedasticity can be assessed by interpreting the results of Box's test.

Summary

Chapter Three described the methodology that was used to conduct this study. Six research questions guide this study to determine if the reading and mathematics NWEA-MAP assessments were adequate measures to predict student achievement on the language arts and mathematics MCT2 state mandated standardized achievement assessment. NWEA-MAP reading and mathematics RIT scores and language arts and mathematics MCT2 scale scores of sixth and eighth grade students who attended a Southern Mississippi School district during the 2011-2012 school term were used for this

study. Only the scores from the sixth and eighth grade students who took all four assessments during the 2011-2012 school term were used. The researcher used a correlational research design employing regression analysis to answer each research question.

CHAPTER IV

RESULTS

Introduction

The purpose of this study was to determine whether reading and mathematics NWEA-MAP assessments are an adequate measure to predict student achievement on the MCT2. Statistical analyses were conducted using SPSS, version 22. Chapter Four begins with descriptive statistics, the results of pre-analysis data screening to test the assumptions of regression analysis, and the results of data analysis to answer research questions following the preliminary analysis. The chapter concludes with a summary of the results.

Descriptive Statistics

Data used in this study include sixth and eighth grade students' language arts and mathematics MCT2 scale scores, reading and mathematics NWEA-MAP RIT scores, and student characteristics of gender, ethnicity, and economically disadvantaged status. Information from 676 sixth grade students and 659 eighth grade students from a Southern Mississippi school district were included in the analyses. Table 17 represents the percentage of students included in the analyses by ethnicity, gender, and economically disadvantaged status. Table 18 represents the total percentage and total number of

students included in the analyses who obtained MCT2 performance levels of Advanced, Proficient, Basic, or Minimum.

Table 17

Student Ethnicity, Gender, and Economically Disadvantaged Status Percentages

Grade	Caucasian	African American	Other	Male	Female	Economically Disadvantaged Yes	Economically Disadvantaged No
6	71.0	22.3	6.7	49.6	50.4	0.4	99.6
8	66.5	26.3	7.3	48.6	51.4	7.9	92.1

Table 18

Percentage and number of sixth and eighth grade students who scored Advanced, Proficient, Basic, or Minimum on the Language Arts and Mathematics MCT2.

Performance Level	Language Arts				Mathematics			
	Grade 6		Grade 8		Grade 6		Grade 8	
	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage
Minimum	31	4.6	18	2.7	44	6.5	17	2.6
Basic	151	22.3	148	22.5	93	13.8	59	9.0
Proficient	380	56.2	434	65.9	338	50.0	383	58.1
Advanced	114	16.9	59	9.0	201	29.7	200	30.3

Results of Question One

In order to answer Research Question One: How accurately do fall NWEA-MAP RIT reading scores predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade, the researcher performed a simple linear regression analysis with students' MCT2 language arts scale scores as the

dependent variable and students' fall reading NWEA-MAP RIT scores, as the independent variable.

A simple linear regression analysis was performed with sixth grade students' MCT2 language arts scale scores as the dependent variable and sixth grade students' fall reading NWEA-MAP RIT scores as the independent variable. Scores from 676 sixth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.820. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified eight outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between sixth grade students' fall reading NWEA-MAP RIT scores and their MCT2 language arts scale scores, $F(1, 674) = 858.893$, $p = .000$. Fall reading NWEA-MAP RIT scores explained 56% of the variability in MCT2 language arts scale scores. Therefore, sixth grade students' fall reading NWEA-MAP RIT scores are a reliable predictor of MCT2 language arts scale scores. The regression equation is: predicted MCT2 language arts scale scores = 15.527 + 0.648 (fall reading NWEA-MAP RIT scores). Tables 19 and 20 detail the sixth grade results of the simple linear regression analysis for Research Question One.

Table 19

Research Question One Coefficients for Sixth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
Constant	15.527	4.756		3.265	.001
Fall NWEA-MAP Reading RIT Score	.648	.022	.749	29.307	.000

Note. Dependent Variable: MCT2 Reading Scale Score.

Table 20

Research Question One ANOVA Table for Sixth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	38256.540	1	38256.540	858.893	.000
Residual	30021.099	674	44.542		
Total	68277.639	675			

Note: Dependent Variable: MCT2 Reading Scale Score. Predictors: (Constant), Fall NWEA-MAP Reading RIT Score.

A simple linear regression analysis was performed with eighth grade students' MCT2 language arts scale scores as the dependent variable and eighth grade students' fall reading NWEA-MAP RIT scores as the independent variable. Scores from 659 eighth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.840. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified seven outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between eighth grade students' fall reading NWEA-MAP RIT

scores and their MCT2 language arts scale scores, $F(1, 657) = 722.560$, $p = .000$. Fall reading NWEA-MAP RIT scores explained 52.3% of the variability in MCT2 language arts scale scores. Therefore, eighth grade students' fall reading NWEA-MAP RIT scores are a reliable predictor of MCT2 language arts scale scores. The regression equation is: predicted MCT2 language arts scale scores = 24.490 + 0.593 (fall reading NWEA-MAP RIT scores). Tables 21 and 22 detail the eighth grade results of the simple linear regression analysis for Research Question One.

Table 21

Research Question One Coefficients for Eighth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
Constant	24.490	4.892		5.006	.000
Fall NWEA-MAP Reading RIT Score	.593	.022	.724	26.880	.000

Note. Dependent Variable: MCT2 Reading Scale Score.

Table 22

Research Question One ANOVA Table for Eighth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	30407.968	1	30407.968	722.560	.000
Residual	27648.964	657	42.084		
Total	58056.932	658			

Note. Dependent Variable: MCT2 Reading Scale Score. Predictors: (Constant), Fall NWEA-MAP Reading RIT Score.

Results of Question Two

In order to answer Research Question Two: How accurately do fall NWEA-MAP RIT mathematics scores predict mathematics achievement scores as measured by the

MCT2 mathematics scale scores for students in sixth and eighth grade, the researcher performed a simple linear regression analysis with students' MCT2 mathematics scale scores as the dependent variable and students' fall mathematics NWEA-MAP RIT scores, as the independent variable.

A simple linear regression analysis was performed with sixth grade students' MCT2 mathematics scale scores as the dependent variable and sixth grade students' fall mathematics NWEA-MAP RIT scores as the independent variable. Scores from 676 sixth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 2.031. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified seven outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no

appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between sixth grade students' fall mathematics NWEA-MAP RIT scores and their MCT2 mathematics scale scores, $F(1, 674) = 1315.280$, $p = .000$. Fall mathematics NWEA-MAP RIT scores explained 66.1% of the variability in MCT2 mathematics scores. Therefore, sixth grade students' fall mathematics NWEA-MAP RIT scores are a reliable predictor of MCT2 mathematics scale scores. The regression equation is: predicted MCT2 mathematics scale scores = 9.287 + 0.672 (fall mathematics NWEA-MAP RIT scores). Tables 23 and 24 detail the sixth grade results of the simple linear regression analysis for Research Question Two.

Table 23

Research Question Two Coefficients for Sixth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
Constant	9.287	4.102		2.264	.024
Fall NWEA-MAP Mathematics RIT Score	.672	.019	.813	36.267	.000

Note. Dependent Variable: MCT2 Mathematics Scale Score.

Table 24

Research Question Two ANOVA Table for Sixth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	30407.968	1	55065.773	1315.280	.000
Residual	27648.964	674	41.866		
Total	58056.932	675			

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Fall NWEA-MAP Mathematics RIT Score.

A simple linear regression analysis was performed with eighth grade students' MCT2 mathematics scale scores as the dependent variable and eighth grade students' fall mathematics NWEA-MAP RIT scores as the independent variable. Scores from 659 eighth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.859. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified three outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples

larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between eighth grade students' fall mathematics NWEA-MAP RIT scores and their MCT2 mathematics scale scores, $F(1, 657) = 1073.604$, $p = .000$. Fall mathematics NWEA-MAP RIT scores explained 62% of the variability in MCT2 mathematics scores. Therefore, eighth grade students' fall mathematics NWEA-MAP RIT scores are a reliable predictor of MCT2 mathematics scale scores. The regression equation is: predicted MCT2 mathematics scale scores = 42.360 + 0.505 (fall mathematics NWEA-MAP RIT scores). Tables 25 and 26 detail the sixth grade results of the simple linear regression analysis for Research Question Two.

Table 25

Research Question Two Coefficients for Eighth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
Constant	42.360	3.577		11.841	.000
Fall NWEA-MAP Mathematics RIT Score	.505	.015	.788	32.766	.000

Note. Dependent Variable: MCT2 Mathematics Scale Score.

Table 26

Research Question Two ANOVA Table for Eighth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	33473.789	1	33473.789	1073.604	.000
Residual	20484.532	657	31.179		
Total	53958.322	658			

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Fall NWEA-MAP Mathematics RIT Score.

Results of Question Three

In order to answer Research Question Three: How accurately do spring NWEA-MAP RIT reading scores predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade, the researcher performed a simple linear regression analysis with students' MCT2 language arts scale scores as the dependent variable and students' spring reading NWEA-MAP RIT scores, as the independent variable.

A simple linear regression analysis was performed with sixth grade students' MCT2 language arts scale scores as the dependent variable and sixth grade students' spring reading NWEA-MAP RIT scores as the independent variable. Scores from 676 sixth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.708. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used

two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified four outliers with standardized residuals greater than ± 3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between sixth grade students' spring reading NWEA-MAP RIT scores and their MCT2 language arts scale scores, $F(1, 674) = 690.437$, $p = .000$. Spring reading NWEA-MAP RIT scores explained 50.5% of the variability in MCT2 language arts scale scores. Therefore, sixth grade students' spring reading NWEA-MAP RIT scores are a reliable predictor of MCT2 language arts scale scores. The regression equation is: predicted MCT2 language arts scale scores = 14.373 + 0.640 (spring reading NWEA-MAP RIT scores). Tables 27 and 28 detail the sixth grade results of the simple linear regression analysis for Research Question Three.

Table 27

Research Question Three Coefficients for Sixth Grade Analysis

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	14.373	5.347		2.688	.007
Spring NWEA-MAP Reading RIT Score	.640	.024	.711	26.276	.000

Note. Dependent Variable: MCT2 Reading Scale Score.

Table 28

Research Question Three ANOVA Table for Sixth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	34550.079	1	34550.079	690.437	.000
Residual	33727.560	674	50.041		
Total	68277.639	675			

Note. Dependent Variable: MCT2 Reading Scale Score. Predictors: (Constant), Spring NWEA-MAP Reading RIT Score.

A simple linear regression analysis was performed with eighth grade students' MCT2 language arts scale scores as the dependent variable and eighth grade students' spring reading NWEA-MAP RIT scores as the independent variable. Scores from 659 eighth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.789. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified five outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples

larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between eighth grade students' spring reading NWEA-MAP RIT scores and their MCT2 language arts scale scores, $F(1, 657) = 723.510$, $p = .000$. Spring reading NWEA-MAP RIT scores explained 52.3% of the variability in MCT2 language arts scale scores. Therefore, eighth grade students' spring reading NWEA-MAP RIT scores are a reliable predictor of MCT2 language arts scale scores. The regression equation is: predicted MCT2 language arts scale scores = 24.745 + 0.583 (spring reading NWEA-MAP RIT scores). Tables 29 and 30 detail the eighth grade results of the simple linear regression analysis for Research Question Three.

Table 29

Research Question Three Coefficients for Eighth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
Constant	24.745	4.880		5.071	.000
Spring NWEA-MAP Reading RIT Score	.583	.022	.724	26.898	.000

Note. Dependent Variable: MCT2 Reading Scale Score.

Table 30

Research Question Three ANOVA Table for Eighth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	30427.004	1	30427.004	723.510	.000
Residual	27629.928	657	42.055		
Total	58056.932	658			

Note. Dependent Variable: MCT2 Reading Scale Score. Predictors: (Constant), Spring NWEA-MAP Reading RIT Score.

Results of Question Four

In order to answer Research Question Four: How accurately do spring NWEA-MAP RIT mathematics scores predict mathematics achievement scores as measured by the MCT2 mathematics scale scores for students in sixth and eighth grade, the researcher performed a simple linear regression analysis with students' MCT2 mathematics scale scores as the dependent variable and students' spring mathematics NWEA-MAP RIT scores, as the independent variable.

A simple linear regression analysis was performed with sixth grade students' MCT2 mathematics scale scores as the dependent variable and sixth grade students' spring mathematics NWEA-MAP RIT scores as the independent variable. Scores from 676 sixth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.916. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual

inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified six outliers with standardized residuals greater than ± 3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between sixth grade students' spring mathematics NWEA-MAP RIT scores and their MCT2 mathematics scale scores, $F(1, 674) = 1115.387$, $p = .000$. Spring mathematics NWEA-MAP RIT scores explained 62.3% of the variability in MCT2 mathematics scale scores. Therefore, sixth grade students' spring mathematics NWEA-MAP RIT scores are a reliable predictor of MCT2 mathematics scale scores. The regression equation is: predicted MCT2 mathematics scale scores = 4.811 + 0.667 (spring mathematics NWEA-MAP RIT scores). Tables 31 and 32 detail the sixth grade results of the simple linear regression analysis for Research Question Four.

Table 31

Research Question Four Coefficients for Sixth Grade Analysis

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	4.811	4.588		1.049	.295
Spring NWEA-MAP Mathematics RIT Score	.667	.020	.790	33.397	.000

Note. Dependent Variable: MCT2 Mathematics Scale Score.

Table 32

Research Question Four ANOVA Table for Sixth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	51913.555	1	51913.555	1115.387	.000
Residual	31370.035	674	46.543		
Total	83283.590	675			

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Spring NWEA-MAP Mathematics RIT Score.

A simple linear regression analysis was performed with eighth grade students' MCT2 mathematics scale scores as the dependent variable and eighth grade students' spring mathematics NWEA-MAP RIT scores as the independent variable. Scores from 659 eighth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. Visual inspection of the scatterplot reveals the dependent variable and independent variable are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.929. To check for homoscedasticity, the researcher plotted the Regression Standardized Residual against the Regression Standardized Predicted value. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used

two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified six outliers with standardized residuals greater than ± 3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. All assumptions have been met; therefore, simple linear regression is an appropriate analysis.

The result of the simple linear regression revealed that there was a statistically significant relationship between eighth grade students' spring mathematics NWEA-MAP RIT scores and their MCT2 mathematics scale scores, $F(1, 657) = 1115.669$, $p = .000$. Spring mathematics NWEA-MAP RIT scores explained 63.8% of the variability in MCT2 mathematics scale scores. Therefore, eighth grade students' spring mathematics NWEA-MAP RIT scores are a reliable predictor of MCT2 mathematics scale scores. The regression equation is: predicted MCT2 mathematics scale scores = 40.504 + 0.504 (spring mathematics NWEA-MAP RIT scores). Tables 33 and 34 detail the eighth grade results of the simple linear regression analysis for Research Question Four.

Table 33

Research Question Four Coefficients for Eighth Grade Analysis

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	40.504	3.498		11.578	.000
Spring NWEA-MAP Mathematics RIT Score	.504	.015	.799	34.039	.000

Note. Dependent Variable: MCT2 Mathematics Scale Score.

Table 34

Research Question Four ANOVA Table for Eighth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	34433.498	1	34433.498	1158.669	.000
Residual	19524.824	657	29.718		
Total	53958.322	658			

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Spring NWEA-MAP Mathematics RIT Score.

Results of Question Five

In order to answer Research Question Five: How accurately do the fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status predict reading achievement scores as measured by the MCT2 language arts scale scores for students in sixth and eighth grade, the researcher performed a multiple regression analysis with students' MCT2 language arts scale scores as the dependent variable and students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, characteristics of gender, ethnicity, and economically disadvantaged status as the independent variables.

When using multiple regression analysis, each categorical variable is limited to having no more than two categories. There were three categories (Caucasian, African American, and Other) to the independent variable, ethnicity; therefore, the researcher dummy coded the variable. When including a categorical variable with more than two categories in multiple regression, the researcher must recode the categorical variable into separate dichotomous variables. For the purposes of Research Question Five, the ethnicity variable has been dummy coded into three separate dichotomous variables: Caucasian, African American, and Other. The researcher identified Caucasian as the comparison variable because the number of participants who were Caucasian were significantly larger than the other two recoded variables.

A standard multiple regression analysis was performed with sixth grade students' MCT2 language arts scale scores as the dependent variable and sixth grade students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status as the independent variables. Scores from 676 sixth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. The researcher also plotted Partial Regression Plots of each independent variable and the dependent variable. Visual inspection of the scatterplots reveals the dependent variable and independent variables are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.717. To check for homoscedasticity, the researcher plotted the Regression Studentized Residuals against the Regression Unstandardized Predicted values. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to

determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified eight outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. Multicollinearity happens when moderate to high intercorrelations occur among independent variables used in a regression analysis. To assess for multicollinearity, the researcher reviewed the tolerance statistics and the variance inflation factor. Tolerance values less than 0.1 and variance inflation factors greater than 10 indicate issues with multicollinearity. Tolerance values for each independent variable are all greater than 0.1; variance inflation factors for each independent variable are all less than 10; therefore, multicollinearity is not a concern. All assumptions have been met; therefore, multiple regression is an appropriate analysis.

A standard multiple regression was conducted to evaluate how well sixth grade students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status predicted reading achievement scores as measured by the MCT2 language arts scale scores. The linear combination of fall reading NWEA-MAP RIT scores, spring reading

NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 language arts scale scores, $F(6, 669) = 171.273$, $p = .000$. The multiple correlation coefficient was .602, indicating that approximately 60.2% of the variance of MCT2 language arts scale scores can be accounted for by the linear combination of students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status. The regression equation for predicting MCT2 language arts scale scores is: predicted language arts MCT2 scale scores = 1.357 - 0.145 (African American) - 0.448 (other) - 1.843 (economically disadvantaged) + 1.638 (gender) + 0.287 spring reading NWEA-MAP RIT scores) + 0.426 (fall reading NWEA-MAP RIT scores). Only gender, fall reading NWEA-MAP scores, and spring reading NWEA-MAP scores contributed statistically significantly to the equation. Tables 35, 36 and 37 detail the sixth grade results of the multiple regression analysis for Research Question Five.

Table 35

Research Question Five Model Summary for Sixth Grade Analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.778	0.606	0.602	6.344

Note. Dependent Variable: MCT2 Language Arts Scale Score. Predictors: (Constant), Spring NWEA-MAP Reading RIT Score, Fall NWEA-MAP Reading RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 36

Research Question Five ANOVA Table for Sixth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	41355.110	6	6892.518	171.23	.000
Residual	26922.529	669	40.243		
Total	68277.639	675			

Note. Dependent Variable: MCT2 Language Arts Scale Score. Predictors: (Constant), Spring NWEA-MAP Reading RIT Score, Fall NWEA-MAP Reading RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 37

Research Question Five Coefficients for Sixth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
1 (Constant)	1.357	8.341		0.163	0.871
African American	-0.145	0.607	-0.006	-0.239	0.811
Other Minority	-0.448	0.989	-0.011	-0.452	0.651
Economically Disadvantaged	-1.843	3.726	-0.012	-0.495	0.621
Gender	1.638	0.493	0.082	3.326	0.001
Fall Reading NWEA-MAP RIT Score	0.426	0.035	0.492	12.318	0.000
Spring Reading NWEA-MAP RIT Score	0.287	0.036	0.320	8.040	0.000

Note. Dependent Variable: MCT2 Language Arts Scale Score.

A standard multiple regression analysis was performed with eighth grade students' MCT2 language arts scale scores as the dependent variable and eighth grade students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status as the independent variables. Scores from 659 eighth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. The researcher also plotted Partial Regression Plots of

each independent variable and the dependent variable. Visual inspection of the scatterplots reveals the dependent variable and independent variables are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.867. To check for homoscedasticity, the researcher plotted the Regression Studentized Residuals against the Regression Unstandardized Predicted values. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified six outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. Multicollinearity happens when moderate to high intercorrelations occur among independent variables used in a regression analysis. To assess for multicollinearity, the researcher reviewed the tolerance statistics and the variance inflation factor. Tolerance values less than 0.1 and variance inflation factors greater than 10 indicate issues with multicollinearity. Tolerance values for each independent variable are all greater than 0.1; variance inflation factors for each independent variable are all less than 10; therefore, multicollinearity is not a concern. All assumptions have been met; therefore, multiple regression is an appropriate analysis.

A standard multiple regression was conducted to evaluate how well eighth grade students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status predicted reading achievement scores as measured by the MCT2 language arts scale scores. The linear combination of fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 language arts scale scores, $F(6, 652) = 172.660$, $p = .000$. The multiple correlation coefficient was .610, indicating that approximately 61% of the variance of MCT2 language arts scale scores can be accounted for by the linear combination of students' fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status. The regression equation for predicting MCT2 language arts scale scores is: predicted language arts MCT2 scale scores = 8.981 – 1.374 (African American) + 3.421 (other) + 1.874 (economically disadvantaged) + 1.061 (gender) + 0.320 spring reading NWEA-MAP RIT scores) + 0.315 (fall reading NWEA-MAP RIT scores). Each of the independent variables contributed statistically significantly to the equation. Tables 38, 39 and 40 detail the eighth grade results of the multiple regression analysis for Research Question Five.

Table 38

Research Question Five Model Summary for Eighth Grade Analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.783	0.614	0.610	5.865

Note. Dependent Variable: MCT2 Language Arts Scale Score. Predictors: (Constant), Spring NWEA-MAP Reading RIT Score, Fall NWEA-MAP Reading RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 39

Research Question Five ANOVA Table for Eighth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	35631.574	6	5938.596	172.660	.000
Residual	22425.357	652	34.395		
Total	58056.932	658			

Note. Dependent Variable: MCT2 Language Arts Scale Score. Predictors: (Constant), Spring NWEA-MAP Reading RIT Score, Fall NWEA-MAP Reading RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 40

Research Question Five Coefficients for Eighth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
1 (Constant)	8.981	4.816		1.865	0.063
Gender	1.061	0.463	0.057	2.294	0.022
Economically Disadvantaged	1.874	0.890	0.054	2.107	0.036
Fall Reading NWEA-MAP RIT Score	0.315	0.031	0.384	10.049	0.000
Spring Reading NWEA-MAP RIT Score	0.320	0.031	0.398	10.374	0.000
African American	-1.374	0.536	-0.064	-2.565	0.011
Other Minority	3.421	0.895	0.095	3.823	0.000

Note. Dependent Variable: MCT2 Language Arts Scale Score.

Results of Question Six

In order to answer Research Question Six: How accurately do the fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status predict mathematics achievement scores as measured by the MCT2 mathematics scale scores for students in sixth and eighth grade, the researcher performed a multiple regression analysis with students' MCT2 mathematics scale scores as the dependent variable and students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, characteristics of gender, ethnicity, and economically disadvantaged status as the independent variables.

When using multiple regression analysis, each categorical variable is limited to having no more than two categories. There were three categories (Caucasian, African American, and Other) to the independent variable, ethnicity; therefore, the researcher dummy coded the variable. When including a categorical variable with more than two categories in multiple regression, the researcher must recode the categorical variable into separate dichotomous variables. For the purposes of Research Question Six, the ethnicity variable has been dummy coded into three separate dichotomous variables: Caucasian, African American, and Other. The researcher identified Caucasian as the comparison variable because the number of participants who were Caucasian were significantly larger than the other two recoded variables.

A standard multiple regression analysis was performed with sixth grade students' MCT2 mathematics scale scores as the dependent variable and sixth grade students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores,

student characteristics of gender, ethnicity, and economically disadvantaged status as the independent variables. Scores from 676 sixth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. The researcher also plotted Partial Regression Plots of each independent variable and the dependent variable. Visual inspection of the scatterplots reveals the dependent variable and independent variables are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.949. To check for homoscedasticity, the researcher plotted the Regression Studentized Residuals against the Regression Unstandardized Predicted values. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified seven outliers with standardized residuals greater than $\pm$3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. Multicollinearity happens when moderate to high intercorrelations occur among independent variables used in a regression analysis. To assess for multicollinearity, the researcher reviewed the tolerance statistics and the variance inflation factor. Tolerance values less than 0.1 and variance

inflation factors greater than 10 indicate issues with multicollinearity. Tolerance values for each independent variable are all greater than 0.1; variance inflation factors for each independent variable are all less than 10; therefore, multicollinearity is not a concern. All assumptions have been met; therefore, multiple regression is an appropriate analysis.

A standard multiple regression was conducted to evaluate how well sixth grade students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status predicted mathematics achievement scores as measured by the MCT2 mathematics scale scores. The linear combination of fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 mathematics scale scores , $F(6, 669) = 270.399$, $p = .000$. The multiple correlation coefficient was .705, indicating that approximately 70.5% of the variance of MCT2 mathematics scale scores can be accounted for by the linear combination of students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status. The regression equation for predicting MCT2 mathematics scale scores is: predicted mathematics MCT2 scale scores = 18.050 - 2.122 (African American) - 0.379 (other) - 11.192 (economically disadvantaged) + 0.736 (gender) + 0.299 spring mathematics NWEA-MAP RIT scores) + 0.421 (fall mathematics NWEA-MAP RIT scores). Only economically disadvantaged status, ethnicity, fall reading NWEA-MAP scores, and spring reading NWEA-MAP scores contributed statistically significantly to the equation. Tables 41, 42 and 43 detail the sixth grade results of the multiple regression analysis for Research Question Six.

Table 41

Research Question Six Model Summary for Sixth Grade Analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.841	0.708	0.705	6.029

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Spring NWEA-MAP Mathematics RIT Score, Fall NWEA-MAP Mathematics RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 42

Research Question Six ANOVA Table for Sixth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	58967.917	6	9827.986	270.399	.000
Residual	24315.674	669	36.646		
Total	83283.590	675			

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Spring NWEA-MAP Mathematics RIT Score, Fall NWEA-MAP Mathematics RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 43

Research Question Six Coefficients for Sixth Grade Analysis

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t	Sig.
1 (Constant)	18.050	7.823		2.307	0.021
Economically Disadvantaged	-11.192	3.523	-0.067	-3.176	0.002
Gender	0.736	0.467	0.033	1.574	0.116
Fall Mathematics NWEA-MAP RIT Score	0.421	0.033	0.509	12.799	0.000
Spring Mathematics NWEA-MAP RIT Score	0.299	0.033	0.353	8.934	0.000
African American	-2.122	0.572	-0.080	-3.707	0.000
Other Minority	-0.379	0.943	-0.009	-0.402	0.688

Note. Dependent Variable: MCT2 Mathematics Scale Score.

A standard multiple regression analysis was performed with eighth grade students' MCT2 mathematics scale scores as the dependent variable and eighth grade students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status as the independent variables. Scores from 659 eighth grade students were used in the analysis. To check for linearity, the researcher plotted the Studentized Residuals against the Unstandardized Predicted values. The researcher also plotted Partial Regression Plots of each independent variable and the dependent variable. Visual inspection of the scatterplots reveals the dependent variable and independent variables are likely linear. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.944. To check for homoscedasticity, the researcher plotted the Regression Studentized Residuals against the Regression Unstandardized Predicted values. Visual inspection of the scatterplot reveals that there is homoscedasticity. The researcher used two methods to determine whether residuals were normally distributed: histogram and Normal P-Plot. Visual inspection of the histogram reveals that the standardized residuals appear to be normally distributed and the Normal P-Plot confirms that the residuals are normally distributed. The researcher reviewed the Casewise Diagnostics to determine if SPSS identified any outliers. SPSS identified four outliers with standardized residuals greater than ± 3 standard deviations. According to Mertler and Vannata (2010), samples larger than 100 are likely to contain a few outliers; however, if the researcher cannot determine if the subject is different from the sample, cases should not be dropped, but rather be included in the analysis. Analyses were conducted with and without the outliers and no appreciable differences were found in the results. Multicollinearity happens when

moderate to high intercorrelations occur among independent variables used in a regression analysis. To assess for multicollinearity, the researcher reviewed the tolerance statistics and the variance inflation factor. Tolerance values less than 0.1 and variance inflation factors greater than 10 indicate issues with multicollinearity. Tolerance values for each independent variable are all greater than 0.1; variance inflation factors for each independent variable are all less than 10; therefore, multicollinearity is not a concern. All assumptions have been met; therefore, multiple regression is an appropriate analysis.

A standard multiple regression was conducted to evaluate how well eighth grade students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status predicted mathematics achievement scores as measured by the MCT2 mathematics scale scores. The linear combination of fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 mathematics scale scores , $F(6, 652) = 241.383$, $p = .000$. The multiple correlation coefficient was .687, indicating that approximately 68.7% of the variance of MCT2 mathematics scale scores can be accounted for by the linear combination of students' fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status. The regression equation for predicting MCT2 mathematics scale scores is: predicted mathematics MCT2 scale scores = 32.969 - 0.765 (African American) + 1.953 (other) - 0.010 (economically disadvantaged) + 0.609 (gender) + 0.290 (spring mathematics NWEA-MAP RIT scores) + 0.247 (fall mathematics NWEA-MAP RIT scores). Only ethnicity, fall reading NWEA-

MAP scores, and spring reading NWEA-MAP scores contributed statistically significantly to the equation. Tables 44, 45 and 46 detail the eighth grade results of the multiple regression analysis for Research Question Six.

Table 44

Research Question Six Model Summary for Eighth Grade Analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.830	0.690	0.687	5.069

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Spring NWEA-MAP Mathematics RIT Score, Fall NWEA-MAP Mathematics RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 45

Research Question Six ANOVA Table for Eighth Grade Analysis

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	37207.932	6	6201.322	241.383	.000
Residual	16750.390	652	25.691		
Total	53958.322	658			

Note. Dependent Variable: MCT2 Mathematics Scale Score. Predictors: (Constant), Spring NWEA-MAP Mathematics RIT Score, Fall NWEA-MAP Mathematics RIT Score, Gender, Ethnicity, Economically Disadvantaged.

Table 46

Research Question Six Coefficients for Eighth Grade Analysis

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
1 (Constant)	32.968	3.587		9.192	0.000
African American	-0.765	0.468	-0.037	-1.635	0.102
Other Minority	1.953	10.776	0.056	2.518	0.012
Economically Disadvantaged	-0.010	0.762	0.000	-0.014	0.989
Gender	0.609	0.399	0.034	1.527	0.127
Spring Mathematics NWEA-MAP RIT Score	0.290	0.026	0.459	11.296	0.000
Fall Mathematics NWEA-MAP RIT Score	0.247	0.026	0.386	9.431	0.000

Note. Dependent Variable: MCT2 Mathematics Scale Score.

Summary

Chapter Four describes the results of the regression analyses used to answer the six research questions that guide this study. Simple linear regression analyses were conducted to answer the first four research questions. Standard multiple regression analyses were conducted to answer the final two research questions. The results of Research Question One indicate that both sixth and eighth grade students' fall reading NWEA-MAP RIT scores are a reliable predictor of MCT2 language arts scale scores. The results of Research Question Two indicate that both sixth and eighth grade students' fall mathematics NWEA-MAP RIT scores are a reliable predictor of MCT2 mathematics scale scores. The results of Research Question Three indicate that both sixth and eighth grade students' spring reading NWEA-MAP RIT scores are a reliable predictor of MCT2 language arts scale scores. The results of Research Question Four indicate that sixth and eighth grade students' spring mathematics NWEA-MAP RIT scores are a reliable predictor of MCT2 mathematics scale scores. The results of Research Question Five

indicate that for sixth and eighth grade students, the linear combination of fall reading NWEA-MAP RIT scores, spring reading NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 language arts scale scores. The results of Research Question Six indicate that for sixth and eighth grade students, the linear combination of fall mathematics NWEA-MAP RIT scores, spring mathematics NWEA-MAP RIT scores, student characteristics of gender, ethnicity, and economically disadvantaged status was significantly related to MCT2 mathematics scale scores. Chapter Five contains the summary, conclusions, and recommendations based on the results of the six research questions.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Introduction

Chapter Five presents quantitative evidence that answers each of the research questions that guided this study. The researcher found evidence that there is a strong relationship between NWEA-MAP reading and MCT2 language arts assessments, as well as, NWEA-MAP mathematics and MCT2 mathematics assessments; therefore, NWEA-MAP assessment scores can be used to predict MCT2 scores. Moreover, the researcher found that the use of NWEA-MAP assessments, as well as other student characteristics, may provide prescriptive and diagnostic information to teachers, principals, parents, school psychologists, and other educators who may use this information to prepare students for successful achievement on state mandated standardized assessments. This chapter provides the summary, conclusions, limitations, recommendations, and implications for future research.

Summary

The purpose of this study was to determine whether sixth and eighth grade students' language arts and mathematics MCT2 achievement scores can be predicted by NWEA-MAP reading and mathematics RIT scores; moreover, the researcher wanted to determine whether student characteristics of gender, economically disadvantaged status,

and ethnicity could be used to predict MCT2 language arts and mathematics performance. Data from 676 sixth grade students and 659 eighth grade students from a Southern Mississippi school district were included in the analyses. Data used in this study included NWEA-MAP reading RIT scores, NEWA-MAP mathematics RIT scores, MCT2 language arts scale scores, and MCT2 mathematics scale scores, and student characteristics of gender, ethnicity, and economically disadvantaged status. A correlational research design was used to answer the six research questions that guide this study. Specifically, the researcher used simple linear regression to answer the first four research questions; whereas, standard multiple regression was used to answer the final two research questions.

Conclusions and Implications

The results of the statistical analysis indicate that fall NWEA-MAP RIT scores predicted the MCT2 scores. Specifically, the fall NWEA-MAP RIT reading scores predicted reading achievement as measured by the MCT2 language arts scale scores and the fall NWEA-MAP RIT mathematics scores predicted math achievement as measured by the MCT2 mathematics scale scores. Shields (2008) also found that NWEA-MAP mathematics RIT scores are a valid predictor for the state-mandated math assessment within the Missouri Assessment Program for students in Grades 6, 7, and 8. The results of the statistical analysis also indicate that spring NWEA-MAP RIT scores predicted the MCT2 scores. Specifically, the spring NWEA-MAP RIT reading scores predicted reading achievement as measured by MCT2 language arts scale scores; moreover, spring NWEA-MAP RIT mathematics scores predicted mathematics achievement as measured by MCT2 mathematics scale scores. Shields (2008) also found that NWEA-MAP reading

RIT scores are a valid predictor for the state-mandated communication arts assessment within the Missouri Assessment Program for students in Grades 6, 7, and 8. Moreover, Andren (2010) found that NWEA-MAP reading RIT scores are a valid predictor for the state-mandated New England Common Assessments Programs. These results indicate that reading and mathematics NWEA-MAP assessments are a valid predictor of state-mandated achievement assessments.

Having said that, schools, districts, and states, have been searching for ways to ensure students are prepared to pass state-mandated assessments. Since MAP assessments have been found to predict performance on those assessments, MAP may be used to help prepare students to pass those assessments. Moreover, because MAP assessments have been shown to identify students at-risk of failing state-mandated assessments and to provide prescriptive and diagnostic information, MAP assessments may also be used for Response to Intervention (RTI) early identification and screening purposes. RTI is an early identification and prevention framework designed to identify at-risk students and provide interventions in a systematic and timely manner (Shapiro & Gebhardt, 2012). Computer-adaptive assessments have emerged as a reliable option for universal screening for RTI purposes (Shapiro & Gebhart, 2012). As the results of this study show, NWEA-MAP assessment results may be used to identify at-risk students for RTI efforts. Additionally, once at-risk students have been identified, NWEA-MAP assessments can be used for their diagnostic and prescriptive information to help guide instructional and intervention practices to remediate at-risk students.

Additionally, this study showed that a linear combination of fall and spring NWEA-MAP RIT reading score, gender, ethnicity and economically disadvantaged

status predicted reading achievement as measured by MCT2 language arts scale scores. Specifically, for sixth grade participants, only gender, fall NWEA-MAP reading RIT score, and spring NWEA-MAP RIT score added significantly to the equation. For eighth grade participants, each of the independent variables added significantly to the equation. Similarly, this study showed that a linear combination of fall and spring NWEA-MAP RIT mathematics score, gender, ethnicity and economically disadvantaged status predicted mathematics achievement as measured by MCT2 mathematics scale scores. Specifically, for sixth grade participants, only economically disadvantaged status, fall NWEA-MAP mathematics RIT score, spring NWEA-MAP mathematics RIT score, and ethnicity added significantly to the equation. For eighth grade students, only ethnicity, spring NWEA-MAP mathematics RIT score, and fall NWEA-MAP mathematics RIT score added statistically significantly to the equation. Similarly, Hall-Michalcewiz (2008), found that reading NWEA-MAP, mathematics NWEA-MAP, ethnicity, and gender were valid predictors of the Delaware State Testing Program mathematics achievement scores.

Limitations

In correlational research, there are threats to internal and external validity. Internal validity refers to the extent to which changes in the dependent variable are directly related to the independent variable and are not due to some other variable; while external validity refers to whether or not the results of research are generalizable to other settings and groups outside of the research setting (Fraenkel et al., 2012). Examples of threats to internal validity include: (a) subject characteristics, (b) mortality, (c) location, (d) instrumentation, (e) data collector characteristics, (f) data collector bias, (g) testing,

(h) history, (i) maturation, (j) regression, and (k) implementation (Fraenkel et al., 2012). Examples of threats to external validity include: population generalizability and ecological generalizability (Fraenkel et al., 2012). It is important for the researcher to be aware of extraneous variables that may explain any results that are obtained through research, as well as factors that may contribute to research results not being generalizable to other populations or settings.

When the researcher finds that two or more characteristics of individuals or groups are correlated there is the possibility that other characteristics can also explain the relationships. When subjects are administered instruments in different specified location, a location threat is possible. In observational studies, the researcher should be cautious of instrument decay or 'observer drift'; ensure that the observers do not become tiered, bored or inattentive. Additionally, if different data collectors administer instruments, data collector characteristics (i.e., gender, age, ethnicity, etc.) may cause a threat. Data collector bias, unconscious bias by the data collector, can also cause threats to validity. A testing threat occurs after a subject has had the experience of responding to the initial instrument that is administered in correlational research. Lastly, mortality is considered a threat to external validity; since the subject must be excluded from the study if scores on both variables cannot be obtained, there is the possibility that the exclusion can increase or decreasing the chances of a relationship occurring (Fraenkel et al., 2012).

The current study does have limitations that should be considered. The following threats to internal validity were noted. First, this study offers findings related to using NWEA-MAP assessments to predict MCT2 performance; however, there are very few studies to report this type of data. Therefore, it is unknown whether the preceding results

would be similar to other findings. Further studies examining predicting MCT2 performance using NWEA-MAP assessments are warranted. Lastly, only students who had all data sets (fall and spring NWEA-MAP RIT mathematics score, fall and spring NWEA-MAP RIT reading score, MCT2 language arts scale score, and MCT2 mathematics scale score) were included in this study. Although, mortality is often unpreventable, the researcher should make an effort to avoid losing participants.

There are threats to external validity that should be considered. The study was conducted in one Mississippi school district. Data from sixth and eighth grade students were used. The scores from this district and those particular grade levels may not be representative of other school districts and same grade levels within Mississippi. The results of this study may also not be generalizable to other states. Replication of the findings across other districts in Mississippi, as well as in other states is warranted.

Recommendations

NCLB sets challenging accountability mandates with the intention of ensuring all students receive a high-quality education as measured by standardized test scores (U.S. Department of Education, 2002a). NCLB requires schools to identify and provide remediation to students at-risk of academic failure (Smith, 2005). States, districts, schools, and educators have made many efforts to ensure students are achieving favorable performance levels on state-mandated assessments. Their efforts include using prescriptive or diagnostic assessments to guide instructional practices. It is important for educators to understand students' strengths and weaknesses, and it is equally important for educators to be able to identify students who are at-risk of failing state-mandated achievement assessments (Smith, 2005). The results of this study indicate that NWEA-

MAP assessments are aligned with MCT2 assessments and NWEA-MAP assessment can predict MCT2 performance; therefore, there are implications to use NWEA-MAP assessment data to provide early intervention screening and intervention strategies to students who are at risk of not obtaining a favorable score on the MCT2.

Moreover, since NWEA-MAP assessments have been proven to predict MCT2 performance, it can be given in the fall, winter, and spring to guide instruction and make continuous decisions about student performance. Unlike MCT2 results, which are only available at the end of a school term, NWEA-MAP assessments are accessible before and during a school term; that means that actionable student data is accessible to educators at various points throughout the school term. This data can be used to prepare students for the MCT2 in May.

Future Research

This study investigated whether NEWA-MAP assessments can be used as valid predictors of MCT2 achievement. Moreover, this study investigated whether knowing student characteristics of gender, ethnicity, and economically disadvantaged status contributes to the prediction of MCT2 achievement. The following are recommendations for future research that may contribute to generalizing the results of this study to different states and grade levels:

1. Replicate a similar study in several different Mississippi schools with a significantly larger population and different grade levels to determine if findings are consistent.
2. Replicate a similar study in other states with a significantly larger population and different grade levels to determine if findings are consistent.

3. Replicate this study in several different Mississippi schools with other Mississippi state tests such as the Subject Area Testing Program, Second Edition (SATP2).
4. Further investigate by comparing the predictive ability of other computer-adaptive assessments, such as STAR, that reportedly provide the same type of information (e.g., aligned with state standards, prescriptive, diagnostic, predictive of state-mandated achievement assessments, etc.) as NWEA-MAP assessments.
5. This study included a small sample of minority students, such as African Americans and other minority races. Further investigate ethnicity using a greater number of minority students to determine if a significant difference exists among students.
6. This study included a small sample of students whose economically disadvantaged status was free/reduced lunch. Further investigate economically disadvantaged status using a greater number of students to determine if a significant difference exists among students.
7. This study suggests that NWEA-MAP assessments are predictive of MCT2 performance. As such, students who are identified as not obtaining a favorable score on the MCT2 may be seen as at-risk. After identifying students who are at-risk using NWEA-MAP assessments, provide targeted interventions, and determine the efficacy of early identification and intervention to improve MCT2 scores.

REFERENCES

Andren, K. (2010). *An analysis of the concurrent and predictive validity of curriculum based measures (CBM), the Measures of Academic Progress (MAP), and the New England Common Assessment Program (NECAP) for reading* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3422255)

Bernhardt, V. (2004). *Data Analysis for Continuous School Improvement (2nd ed.).* New York: Eye on Education.

Buchsbaum, M. (2013). *Logitudinal growth of academic achievement among subgroups using NWEA's MAP* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3562114)

Bush, G. H. W. (2009). *State of the Union Addresses of George H. W. Bush*. Melbourne: Book Jungle.

Carpenter, D. & Ramirez, A. (2012). Exploring the underachievement gap. *Christian Higher Education, 11*(2), 115-130.

Chapman, L.H. (2007, September/October). An update on No Child Left Behind and national trends in education. *Arts Education Policy Review, 109*(1), 25-36.

Delong, S. E. (2007). *A study of the relationship between the utilization of NWEA MAP testing and student achievement* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3286443)

Edwards, A., & Pula, J. (2011). Back to high school: A teacher educator's hands-on encounter with the pressures of high-stakes testing. *Competition and Competitiveness*, 11-14.

Flynn, J. (2008). No Child Left Behind gets a tip from tort theory: Protecting responsible schools against underserved sanctions. *Thomas Jefferson Law Review*, *31*(1), 157-189.

Fraenkel, J., Wallen, N., & Hyun, H. (2012). *How to design and evaluate research in education*. New York: McGraw Hill.

Gibb, S., Fergusson, D. & Horwood, J. (2008). Effects of single-sex and coeducational schooling on the gender gap in educational achievement. *Australian Journal of Education*, *52*(3), 301-317.

Hall-Michalcewiz, J. (2008). *Predicting reading and mathematics Delaware State Testing Program performance using the Measures of Academic Progress* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3312908)

Harriman, N.E. (2005). Perceptions of students and educators on the impact of No Child Left Behind: Some will and some won't. *Rural Special Education Quarterly, 24*(1), 64-69.

Henly, J., McBride, J., Milligan, J., & Nichols, J. (2007). Robbing elementary students of their childhood: The perils of No Child Left Behind. *Education, 128*(1), 56-63.

Lee, J. (2007). Do national state assessments converge for educational accountability? A meta-analytic synthesis of multiple measures in Maine and Kentucky. *Applied Measurement in Education, 20*(2), 171-203.

Maleyko, G., & Gawlik, M. (2011). No Child Left Behind: What we know and what we need to know. *Education*, *131*(3), 600-624.

McNeil, M. (2011). New leeway on horizon under NCLB. *Education Week, 30*(35), 1-12.

Merino, K. & Beckman, T.O. (2010). Using reading curriculum-based measurements as predictors for the Measure of Academic Progress (MAP) standardized test in Nebraska. *International Journal of Psychology: A Biopsychosocial Approach*, 6, 85-98.

Mississippi Department of Education. (2007, April). *Performance level descriptors and item specifications.* Retrieved March 6, 2012 from http://www.mde.k12.ms.us/acad/osa/pld/Performance_Level_Descriptors_and_Test_Item_Specifications_Power_Point.pdf

Mississippi Department of Education. (2010, February). *Understanding the Mississippi statewide accountability system*. Retrieved March 6, 2012, from http://orshome.mde.k12.ms.us/ors/accountability/2009/MSAS-U.pdf

Mississippi Department of Education. (2011). *MCT2 Interpretive Guide*. Retrieved on May 21, 2012 from http://www.mde.k12.ms.us/docs/student-assessment/2011_ig.pdf?sfvrsn=1

Mississippi Department of Education. (2012a). *Interpretive Guide*. Retrieved on April 6, 2013 from https://districtaccess.mde.k12.ms.us/studentassessment/Public%20Access/Forms/AllItems.aspx?RootFolder=%252Fstudentassessment%252FPublic%20Access%252FStatewide_Assessment_Programs%252FAncillaries%252FMCT2

Mississippi Department of Education. (2012b). *Mississippi Curriculum Test, Second Edition (MCT2)*. Retrieved on May 21, 2012 from http://www.mde.k12.ms.us/student-assessment/student-assessment-mct2

Mississippi Department of Education. (2012c). *Mississippi's differentiated accountability model: How the flexibility request impacts school accountability*. Retrieved on April 7, 2013 from http://www.mde.k12.ms.us/docs/federal-programs/esea-flexibility-differentiated-accountability-regional-training-september-2012.pdf?sfvrsn=0

Mississippi Department of Education. (2012d). *Mississippi Public School Accountability Standards 2012*. Retrieved on December 29, 2012 from http://www.mde.k12.ms.us/docs/accreditation-library/revised-10-9-12-2012-stds.pdf?sfvrsn=2

Mississippi Department of Education. (2013a). *Common Core State Standards*. Retrieved on June 1, 2013 from http://www.mde.k12.ms.us/curriculum-and-instruction/curriculum-and-instruction-other-links/common-core-state-standards#resources

Mississippi Department of Education. (2013b). *MSReportCard.com*. Retrieved on December 29, 2012 from http://www.msreportcard.com/rc2012/

National Association of School Psychologists. (n.d.) *What is a school psychologist*? Retrieved from http://www.nasponline.org/resources/handouts/What%20is%20a%20School%20Psych.pdf

National Center for Educational Statistics. (2012). *National assessment of educational progress (NAEP)*. Retrieved May 21, 2012 from http://nces.ed.gov/nationsreportcard/

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Retrieved June 18, 2012 from http://teachertenure.procon.org/sourcefiles/a-nation-at-risk-tenure-april-1983.pdf

National Education Goals Panel. (1993). *The National Education Goals report: Building a nation of learners*. Retrieved June 20, 2012 from http://www2.ed.gov/pubs/goals/report/goalsrpt.txt

National Education Goals Panel. (1999). *The National Education Goals report: Building a nation of learners*. Retrieved June 20, 2012 from http://govinfo.library.unt.edu/negp/reports/99rpt.pdf

National Education Goals Panel. (2002). *History, 1989 to present*. Retrieved June 22, 2012 from http://govinfo.library.unt.edu/negp/page1-7.htm

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards*. National Governors Association Center for Best Practices, Council of Chief State School Officers: Washington D.C.

Neill, M. (2006). Preparing teachers to beat the agonies of NCLB. *Education Digest*, *71*(8), 8-12.

Nishishiba, M., Jones, M., & Kraner, M. (2013). *Research methods and statistics for public and nonprofit administrators: A practical guide*. Los Angeles, CA. Sage Publications, Inc.

Northwest Evaluation Association. (2011). *A study of the alignment of the NWEA RIT Scale with the Mississippi Curriculum Test (MCT)*. Retrieved March 6, 2012 from http://www.nwea.org/sites/www.nwea.org/files/resources/MS_Linking%20Study.pdf

Northwest Evaluation Association. (2012a). *Our History*. Retrieved on March6, 2012 from http://www.nwea.org/about-nwea/our-history

Northwest Evaluation Association. (2012b). *State Information*. Retrieved on March 6, 2012 from http://www.nwea.org/our-research/state-information

Northwest Evaluation Association. (2012c). *The adaptive nature of MAP tests*. Retrieved on March 6, 2012 from http://www.nwea.org/support/article/527/adaptive-nature-map-tests

Northwest Evaluation Association. (2012d). *Measures of academic progress*. Retrieved on March 6 from http://www.nwea.org/products-services/assessments/map%C2%AE

Northwest Evaluation Association. (2012e). *FAQ: Measures of academic progress (MAP)*. Retrieved on May 21, 2012 from http://www.nwea.org/about-nwea/faq/Measures%20of%20Academic%20Progress%20(MAP)#faq-192

Northwest Evaluation Association. (2012f). *The RIT scale*. Retrieved on May 21, 2012 from http://www.nwea.org/support/article/532

Northwest Evaluation Association. (2012g). *Linking MAP to state tests: Proficiency cut score estimation procedures*. Retrieved on January 2, 2013 from http://www.nwea.org/sites/www.nwea.org/files/NWEA%20State%20Standards%20Alignment%20Study%20Methods%20_3_.pdf

Northwest Evaluation Association. (2012h). *The reliability of NWEA results.* Retrieved on February 19, 2013 from http://www.nwea.org/support/article/533

Northwest Evaluation Association. (2014). *Normative data overview*. Retrieved on March 16, 2014 from http://www.nwea.org/sites/www.nwea.org/files/resources/2011_Normative_Data_Overview.pdf

Pearson. (2008, December). *Mississippi Curriculum Test, Second Edition Technical Guide*. Retrieved May 22, 2012 from https://districtaccess.mde.k12.ms.us/studentassessment/Public%20Access/Statewide_Assessment_Programs/-MCT2/MCT2%202007-2008%20Tech%20Manual.pdf

Pearson. (2011, September). *Mississippi Curriculum Test, Second Edition Technical Update*. Retrieved April 6, 2013 from https://districtaccess.mde.k12.ms.us/studentassessment/Public%20Access/Forms/AllItems.aspx?RootFolder=%2Fstudentassessment%2FPublic%20Access%2FStatewide%5FAssessment%5FPrograms%2F%2DMCT2&InitialTabId=Ribbon%2EDocument&VisibilityContext=WSSTabPersistence

Pederson, P.V. (2007, July/August). What is measured is treasured: The impact of the No Child Left Behind Act on non-assessed subjects. *The Clearing House, 80*(6), 287-291.

Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Renaissance Learning. (2011). *The foundation of the STAR assessments: The science of STAR*. Wisconsin Rapids, WI: Renaissance Learning. Retrieved November 10, 2013 from http://doc.renlearn.com/KMNet/R003957507GG2170.pdf

Shapiro, E., & & Gebhardt, S. (2012). Comparing computer-adaptive and curriculum-based measurement methods of assessment. *School Psychology Review, 41*(3), 295-305.

Shields, J. (2008). *A comparison of the NWEA Measures of Academic Progress and the Missouri Assessment Program* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3371096)

Sloane, F. & Kelly, A. (2003). Issues in high-stakes testing programs. *Theory into Practice, 42*(1), 12-17.

Smith, E. & Gorard, S. (2007, June). Improving teacher quality: lessons from America's No Child Left Behind. *Cambridge Journal of Education, 37*(2), 191-206.

Smyth, T. S. (2008, January/February). Who is No Child Left Behind leaving behind? *The Clearing House, 81*(3), 133-137.

Spohn, C. (2008, March/April). Teacher perspectives on No Child Left Behind and arts education: A case study. *Arts Education Policy Review, 109*(4), 3-11.

Tajalli, H. & Opheim, C. (2004). Strategies for closing the gap: Predicting student performance in economically disadvantaged schools. *Educational Research Quarterly*, *28*(4), 44-54.

United States Census Bureau. (n.d.) *American FactFinder: Harrison County*. Retrieved March 6, 2012 from http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_DP_DPDP1

U.S. Department of Education. (1998). *Goals 2000: Reforming education to improve student achievement*. Retrieved June 22, 2012 from http://www2.ed.gov/PDFDocs/g2kfinal.pdf

U.S. Department of Education. (2002a). Public Law 107-110, the *No Child Left Behind Act of 2001*. Retrieved March 6, 2012, from http://www2.ed.gov/policy/elsec/leg/esea02/index.html

U.S. Department of Education. (2002b). *Student achievement and school accountability conference*. Retrieved July 4, 2012 from http://www2.ed.gov/programs/titleiparta/tim/edlite-slide009.html

U.S. Department of Education. (2011). *Obama administration sets high bar for flexibility from No Child Left Behind in order to advance equity and support reform*. Retrieved March 6, 2012, from http://www.ed.gov/news/press-releases/obama-administration-sets-high-bar-flexibility-no-child-left-behind-order-advanc

Wauters, K., Desmet, P. Van den Noortgate, W. (2010). Adaptive item-based learning environments based on the item response theory: Possibilities and challenges. *Journal of Computer Assisted Learning, 26*, 549-562.

APPENDIX A

MISSISSIPPI STATE UNIVERSITY INSTITUTIONAL REVIEW BOARD

APPROVAL

Study 13-347: Predicting Mississippi Curriculum Testing Program, Second Edition Performance Using the Northwest Evaluation Association measures of Academic Progress

3 messages

nmorse@orc.msstate.edu <nmorse@orc.msstate.edu> Mon, Nov 18, 2013 at 1:27 PM
To: mbush25@gmail.com
Cc: nmorse@orc.msstate.edu, IRB@research.msstate.edu, cdh@colled.msstate.edu

November 18, 2013

Mary Cole-Bush
Department of Counseling & Educational Psychology

RE: HRPP Study #13-347: Predicting Mississippi Curriculum Testing Program, Second Edition Performance Using the Northwest Evaluation Association measures of Academic Progress

Dear Ms. Cole-Bush:

This email serves as official documentation that the above referenced project was reviewed and approved via administrative review on 11/18/2013 in accordance with 45 CFR 46.101(b)(4). Continuing review is not necessary for this project. However, in accordance with SOP 01-03 Administrative Review of Applications, a new application must be submitted if the study is ongoing after 5 years from the date of approval. Additionally, any modification to the project must be reviewed and approved by the HRPP prior to implementation. Any failure to adhere to the approved protocol could result in suspension or termination of your project. The HRPP reserves the right, at anytime during the project period, to observe you and the additional researchers on this project.

Please refer to your HRPP number (#13-347) when contacting our office regarding this application.

Thank you for your cooperation and good luck to you in conducting this research project. If you have questions or concerns, please contact me at nmorse@orc.msstate.edu or call 662-325-5220.

Finally, we would greatly appreciate your feedback on the HRPP approval process. Please take a few minutes to complete our survey at http://www.surveymonkey.com/s/YZC7QQD.

Sincerely,

Nicole Morse, CIP
IRB Compliance Administrator

cc: Advisor: Carlen Henington

APPENDIX B

NORTHWEST EVALUATION ASSOCIATION (NWEA) MEASURES OF ACADEMIC PROGRESS (MAP) TABLES AND EXPLANATIONS

Table B1

Estimated Probability of Scoring as Proficient or Higher on the MCT2 Mathematics, by Student Grade Based on Spring RIT Score Range on MAP Mathematics

Grade 8 Estimated Probability (Percentage) of Passing State Test Based on Observed MAP Score			
RIT Range	Probability %	RIT Range	Probability %
120	0	215	29
125	0	220	40
130	0	225	52
135	0	230	65
140	0	235	75
145	0	240	83
150	0	245	89
155	0	250	93
160	0	255	96
165	0	260	97
170	0	265	98
175	1	270	99
180	1	275	99
185	2	280	100
190	3	285	100
195	5	290	100
200	8	295	100
205	13	300	100
210	20		

Note. This table provides the estimated probability of passing the state test based on a MAP test score taken during the spring season.

Table B2

Estimated Probability of Scoring as Proficient or Higher on the MCT2 Language Arts, by Student Grade Based on Spring RIT Score Range on MAP Reading

Grade 8 Estimated Probability (Percentage) of Passing State Test Based on Observed MAP Score			
RIT Range	Probability %	RIT Range	Probability %
120	0	215	33
125	0	220	45
130	0	225	57
135	0	230	69
140	0	235	79
145	0	240	86
150	0	245	91
155	0	250	94
160	0	255	96
165	0	260	98
170	1	265	99
175	1	270	99
180	1	275	100
185	2	280	100
190	4	285	100
195	6	290	100
200	10	295	100
205	15	300	100
210	23		

Note. This table provides the estimated probability of passing the state test based on a MAP test score taken during the spring season.

Table B3

Estimated Probability of Scoring as Proficient or Higher on MCT2 Mathematics, by Student Grade Based on Fall RIT Score Range on MAP Mathematics

Grade 8 Estimated Probability (Percentage) of Passing State Test Based on Observed MAP Score			
RIT Range	Probability %	RIT Range	Probability %
120	0	215	38
125	0	220	50
130	0	225	62
135	0	230	73
140	0	235	82
145	0	240	88
150	0	245	92
155	0	250	95
160	0	255	97
165	0	260	98
170	1	265	99
175	1	270	99
180	2	275	100
185	3	280	100
190	5	285	100
195	8	290	100
200	12	295	100
205	18	300	100
210	27		

Note. This table provides the estimated probability of passing the state test based on a MAP test score taken during the fall season.

Table B4

Estimated Probability of Scoring as Proficient or Higher on the MCT2 Language Arts, by Student Grade Based on Fall RIT Score Range on MAP Reading

Grade 8 Estimated Probability (Percentage) of Passing State Test Based on Observed MAP Score			
RIT Range	Probability %	RIT Range	Probability %
120	0	215	40
125	0	220	52
130	0	225	65
135	0	230	75
140	0	235	83
145	0	240	89
150	0	245	93
155	0	250	96
160	0	255	97
165	0	260	98
170	1	265	99
175	1	270	100
180	2	275	100
185	3	280	100
190	5	285	100
195	8	290	100
200	13	295	100
205	20	300	100
210	29		

Note. This table provides the estimated probability of passing the state test based on a MAP test score taken during the fall season.

Table B5

Percentage of Students whose Pass Status was Accurately Predicted by their MAP Performance Using Reported Cut Scores

Grade 8	MAP Mathematics	MAP Reading
Sample Size	3584	3584
MAP Accurately Predicted State Performance	82.5%	80.4%
MAP Underestimated State Performance	9.1%	10.6%
MAP Overestimated State Performance	8.4%	9.0%

Note. The row labeled MAP Accurately Predicted State Performance shows the percentage of students whose Pass/Not Pass status was predicted accurately when their state test score was linked to their MAP score based on the February 2011 Scale Alignment Study. The row labeled MAP Underestimated State Performance shows the percentage of students whose MAP score predicted they would not pass the state benchmark, but they did pass. The row labeled MAP Overestimated State Performance shows the percentage of students whose MAP score predicted they would pass the state benchmark, but they did not.

APPENDIX C

DEFINITION OF TERMS

Accountability Status (Accountability Label or Accountability Rating): The annual designation assigned to a school or district based on achievement, growth, and if appropriate, high school completion (MDE, 2010).

Achievement Model: The overall school or district level academic performance during the previous school year (MDE, 2010).

Achievement Standards: States must adhere to the Title I requirement to develop academic achievement standards (student performance standards). Student performance standards were developed for each component of the assessment system (MDE, 2010).

Adequate Yearly Progress: Schools, districts, and states are held accountable for student performance under Title I of NCLB based on adequate yearly progress (AYP). State assessments must be the primary indicator in a state's measure of AYP, along with at least one other academic indicator of school performance (U.S. Department of Education, 2002a).

Computerized Adaptive Tests: Computerized adaptive tests are taken on a computer; they are a method for administering tests that adapts to a student's ability level (NWEA, 2012c).

Criterion-Referenced Test (CRT): CRTs compare a student's performance to a specific learning objective or performance standard and not to the performance of other students who have taken the same test (Bernhardt, 2004).

Growth Model: The degree to which a school or district met its expected performance during the previous school year (MDE, 2010).

High-Stakes Testing: High-stakes testing refers to the process of using a test as the primary or only determiner for decisions having major consequences (Fraenkel, Wallen, Hyun, 2012).

Item Response Theory (IRT): The psychometric methodology that allows scores to be computed across different sets of items (NWEA, 2012c).

Measures of Academic Progress (MAP): Computer-based assessment that provides data about what a student knows and what they are ready to learn. MAP is adaptive; it adapts to each response as they take the test. If a student answers a question correctly, the test presents a more challenging item; if a student answers a question incorrectly, MAP offers a simpler item. MAP is aligned to national and Mississippi standards. Currently, NWEA offers MAP for Primary grades (reading and mathematics), MAP (reading, mathematics, language, and science) (NWEA, 2012d).

Mississippi Alternate Assessment of the Extended Curriculum Frameworks (MAAECF): A set of assessments designed for students with disabilities who cannot take the regular statewide assessment even with allowable accommodations and modifications. According to Federal law, the MAAECF can be administered to students with significant cognitive disabilities. The MAAECFF includes assessments in language arts, mathematics, and science. There are several levels of the assessments with separate scoring tables for students in each grade (3-8 and high school). MAAECF results are reported only as proficiency levels (MDE, 2010).

Mississippi Assessment and Accountability Reporting System (MAARS): Online reporting system available to the public that reports annual accountability results. Users

are able to search for assessment results on the state, district, or school level (MDE, 2010).

Mississippi Curriculum Test, Second Edition (MCT2): The MCT2 is based on the revised statewide language arts and mathematics curricula. Tests in language arts and mathematics are administered each year in May to students enrolled in grades three through eight. The results include a numeric scale score and a proficiency level. The proficiency levels represent standards based on cut scores established by committees of Mississippi teachers and approved by the State Board of Education. The proficiency levels are Advanced, Proficient, Basic, and Minimal (MDE, 2010).

National Association of Educational Progress (NAEP): NAEP is the largest nationally representative and continuing assessment of what America's students know and can do in various subject areas (NCES, 2012).

No Child Left Behind (NCLB): NCLB is the reauthorization of the Elementary and Secondary Education Act of 1965 (Bernhardt, 2004).

Norm-Referenced Test (NRT): NRTs are standardized tests. NRTs are designed to measure a broad spectrum of information and to compare the test performance of a school, group, or individual student with the performance of a particular norming group (Bernhardt, 2004).

Northwest Evaluation Association (NWEA): Northwest Evaluation Association is a non-profit organization that was founded in 1974. They are known for being one of the first organizations to create computerized adaptive assessments (NWEA, 2012a).

Performance Level Descriptors: NCLB requires that performance level descriptions for at least three levels, basic, proficient, and advanced are developed. The

performance level descriptors guide the development of the assessments, cut score standard setting, and reporting descriptors and guide teachers' instructional efforts to ensure that students reach proficient levels of performance on the content standards (MDE, 2007).

Quality Distribution Index (QDI): A value that is calculated using data from the MCT2 language arts and mathematics test along with the results of the language arts and mathematics section of the MAAECF. For SATP2, data from Algebra I, Biology I, English II and U.S. History tests along with the results of the language arts and mathematics section of the MAAECF. QDI values range from 0 (100% of students scoring in the lowest proficiency level on the assessments) to 300 (100% of the students scoring in the highest proficiency level on the assessments) (MDE, 2010).

RIT (Rasch Unit): NWEA uses the RIT scale to measure a student's progress. The RIT scale is an equal-interval scale based on the Item Response Theory (NWEA, 2012e).

Scientifically Research-Based: Refers to research where rigorous, systematic and objective guidelines are utilized to obtain reliable and valid results or knowledge pertaining to education activities and programs (Bernhardt, 2004).

Standardized Test: Refers to tests or assessments that have uniformity in content, administration, and scoring (Bernhardt, 2004).